Markovian linearization of random walks on groups

Charles Bordenave, Bastien Dubail

Markovian linearization of random walks on groups

Charles Bordenave* Bastien Dubail†

December 6, 2020

Abstract

In operator algebra, the linearization trick is a technique that reduces the study of a non-commutative polynomial evaluated at elements of an algebra $\mathcal{A}$ to the study of a polynomial of degree one, evaluated on the enlarged algebra $\mathcal{A} \otimes M_r(\mathbb{C})$, for some integer r. We introduce a new instance of the linearization trick which is tailored to study a finitely supported random walk on a group G by studying instead a nearest-neighbor colored random walk on $G \times \{1, \ldots, r\}$, which is much simpler to analyze. As an application we extend well-known results for nearest-neighbor walks on free groups and free products of finite groups to colored random walks, thus showing how one can obtain explicit formulas for the drift and entropy of a finitely supported random walk.

1 Introduction

Let $\mathcal{A}$ be a complex unital algebra and consider a non-commutative polynomial $\mathcal{P} = \mathcal{P}(x_1, \ldots, x_n)$ in the variables $x_1, \ldots, x_n \in \mathcal{A}$. In many cases, a detailed study of the relevant properties of $\mathcal{P}$ is possible only when the degree of $\mathcal{P}$ is small, typically when $\mathcal{P}$ is of degree one, in which case $\mathcal{P}$ is just a linear combination of the x_i's. The linearization trick precisely consists in constructing another polynomial $\tilde{\mathcal{P}}$ of degree one, which can be related to the relevant properties of $\mathcal{P}$. It thus can be used to make computations that were not a priori possible for $\mathcal{P}$. The price to pay is an enlargement of the algebra: writing $\mathcal{P}$ as a finite sum of monomials

$$\mathcal{P} = \sum_{i_1, \ldots, i_k} \alpha_{i_1 \ldots i_k} \, x_{i_1} \cdots x_{i_k},$$

with $\alpha_{i_1 \ldots i_k} \in \mathbb{C}$, $\tilde{\mathcal{P}}$ is generally constructed as

$$\tilde{\mathcal{P}} = \sum_i \tilde{\alpha}_i \otimes x_i$$

*Institut de Mathématiques de Marseille, CNRS & Aix-Marseille University, Marseille, France.
†Département d'Informatique de l'ENS, ENS, CNRS, PSL University, France and INRIA, Paris, France.

where $\tilde{\alpha}_i \in M_r(\mathbb{C})$ are complex matrices. Therefore $\tilde{\mathcal{P}}$ is no longer in the algebra $\mathcal{A}$ but in the algebra $M_r(\mathbb{C}) \otimes \mathcal{A}$ for some integer r that depends on $\mathcal{P}$ (usually through its degree).

Under various names, the linearization trick has been used in several domains such as electrical engineering, random matrices, operators algebras, automata theory, etc, we refer to the survey [13]. A famous application of the linearization trick in operator algebra is the computation of the spectrum of a non-commutative polynomial. If $a \in \mathcal{A}$, the spectrum of a is defined as $\sigma(a) := \{z \in \mathbb{C}, zI - a \text{ is not invertible}\}$. Given a non-commutative polynomial $\mathcal{P}(x_1, \ldots, x_n) \in \mathcal{A}$, it is possible to construct a linear $\tilde{\mathcal{P}} \in M_r(\mathbb{C}) \otimes \mathcal{A}$ for some integer r, such that $(zI - \mathcal{P})$ is invertible in $\mathcal{A}$ if and only if $(\Lambda(z) \otimes I_{\mathcal{A}} - \tilde{\mathcal{P}})$ is invertible in $M_r(\mathbb{C}) \otimes \mathcal{A}$, where $\Lambda(z)$ is the $r \times r$ matrix with only zero elements except the $(1,1)$ entry which is z. Moreover, $(zI - \mathcal{P})^{-1}$ is then precisely the $(1,1)$ entry of $(\Lambda(z) \otimes I_{\mathcal{A}} - \tilde{\mathcal{P}})^{-1}$ (seen as a $r \times r$ matrix of elements in $\mathcal{A}$). Such a construction can be found for instance in the monograph [22, Chapter 10].

As illustrated in this last example, we note that the relevant properties of $\mathcal{P}$ (here its spectrum and its resolvent) dictates the linearization procedure. In this paper, we introduce a new linearization of the Markov kernel of a random walk with finite range on a group. This linearization is a new Markov kernel on an enlarged state space and it corresponds to a nearest-neighbor random walk on the group. Classical quantities such as speed or entropy rate of the original random walk can be read on the new random walk. We illustrate our method on free groups and free product of finite groups. Notably, we establish new formulas for drift and entropy of finite range random walks on these groups.

1.1 The linearization trick and colored random walks on groups

We let G be a finitely generated group with identity element e and we fix a finite set of generators S. We assume that S is symmetric, meaning that $g \in S$ implies $g^{-1} \in S$. Let $p = (p_g)_{g \in G}$ be a probability measure on G. Consider the Markov chain $(X_n)_{n \geq 0}$ on G with transition probabilities

$$\mathbb{P}[X_n = h \mid X_{n-1} = g] = p_{g^{-1}h},$$

for all $g, h \in G$. Such a Markov chain is called a convolution random walk.

The random walk is said to be finitely supported, or to have finite range, if the measure p is finitely supported. It is a nearest-neighbor random walk if p is supported on the set of generators S.

As for any Markov chain, the transition kernel of a right convolutional random walk can be seen as an operator acting on square integrable functions: $f = f(x) \mapsto (\mathcal{P}f)(x) = \sum_y P(x,y) f(y)$. In the present case, it can be written in terms of the left multiplication operators $\lambda(g)$, $g \in G$, defined as: for all $f \in \ell^2(G)$,

$$\lambda(g) \cdot f : x \mapsto f(gx).$$

Letting $\mathcal{P}(g,h) = p_{g^{-1}h}$ it is then possible to write

$$\mathcal{P} = \sum_{g \in G} p_g \lambda(g). \tag{1}$$

This sum is finite if and only if the random walk is finitely supported. Furthermore, λ is a group morphism so $\lambda(gh) = \lambda(g)\lambda(h)$ for all $g, h \in G$ (λ is the left regular representation). Writing each element g in the support of the walk as a product of elements of S, it is thus possible to write $\mathcal{P}$ as a non-commutative polynomial in the operators $\lambda(g), g \in S$. In other words, $\mathcal{P}$ is an element of the left group algebra which is generated by the $\lambda(g)$'s, $g \in G$. The polynomial $\mathcal{P}$ is of degree 1 if and only if the random walk is nearest neighbor. With this point of view, the linearization trick potentially allows to study any finitely supported random walk by considering instead a "nearest-neighbor operator"

$$\tilde{\mathcal{P}} = \sum_{g \in G} \tilde{p}_g \otimes \lambda(g) \tag{2}$$

for some well-chosen matrices $\tilde{p}_g \in M_r(\mathbb{C})$ such that $\sum_g \tilde{p}_g$ is a stochastic matrix, that is, the transition kernel of a Markov chain on $[r]$ where for $n \geq 1$ integer, we set $[n] = \{1, \dots, n\}$.

Then, it turns out that this operator $\tilde{\mathcal{P}}$ is the transition kernel of a Markov chain $(\tilde{X}_n)_n$ on the state space $G \times [r]$, which at each step moves from a pair (g,u) to a pair (h,v) with probability $\tilde{p}_{g^{-1}h}(u,v)$. Equivalently, if one interprets $[r]$ as a set of colors, $\tilde{X}_n$ can at each step be multiplied by $g \in G$ while the color is updated from u to v with probability $p_g(u,v)$. Such a Markov chain will be called a colored random walk and can thus be defined similarly to a classical convolution random walk.

Definition 1. *Let $p = (p_g)_{g \in G}$ be a family of matrices with non-negative coefficients of $M_r(\mathbb{R})$ such that*

$$P = \sum_{g \in G} p_g \tag{3}$$

is a stochastic matrix. A colored random walk is a Markov chain $(Y_n)_n$ on the state space $G \times [r]$, with transition probabilities

$$\mathbb{P}[Y_n = (h,v) \mid Y_{n-1} = (g,u)] = p_{g^{-1}h}(u,v).$$

The support of the random walk is the set of elements $g \in G$ such that $p_g \neq 0$. The colored walk is finitely supported if the support is finite, nearest-neighbor if the support is included in S.

By looking only at the color coordinate, (Y_n) induces a Markov chain on the set of colors $[r]$ whose transition matrix is exactly the matrix P in (3).

Irreducibility assumptions. For standard convolution random walks, it is generally assumed that the support generates the whole group, which then makes the random walk irreducible. For colored random walks, we will make the same assumption but suppose as well that the matrix P defines an irreducible Markov chain on the colors. Note this assumption does not necessarily make the colored walk irreducible, for there may be unreachable pairs (g,u).

Definition 2. *A colored walk is quasi-irreducible if the marginal of its support on G generates the whole group and the matrix P is irreducible.*

Reversibility. An important property satisfied by some Markov chains is reversibility. Recall that, if X is a countable set, ν a measure on X and Q is a Markov chain, then Q is said to be reversible with respect to ν if for all $x,y \in X$, $\nu(x)Q(x,y) = \nu(y)Q(y,x)$. In other words, Q seen as an operator on $\ell^2(X,\nu)$ is self-adjoint. In our setting, $\mathcal{P}$ defined in (1) is said to be reversible if it is reversible for the counting measure on G. This is equivalent to the condition: for all $g \in G$, $p_g = p_{g^{-1}}$. Similarly, consider a colored random walk $\tilde{\mathcal{P}}$ of the form (2). Assume that P defined by (3) has invariant measure π on $[r]$. Then $\tilde{\mathcal{P}}$ is reversible if it is reversible for the product of π and the counting measure on G. This is equivalent to the condition, for all $g \in G$, $u,v \in [r]$,

$$\pi(u)p_g(u,v) = \pi(v)p_{g^{-1}}(v,u).$$

1.2 Main results

Linearizing random walks. The main contribution of this paper is to formalize a way to apply the linearization trick to random walks on groups.

Definition 3. *Let G be a a group generated by a finite set S, with identity element e. Let $(X_n)_{n\geq 0}$ be a random walk with kernel $\mathcal{P}$ as in (1) and with finite support generating G. Let $(Y_n)_{n\geq 0}$ on $G\times[r]$ be a quasi-irreducible, nearest-neighbor colored random walk with kernel $\tilde{\mathcal{P}}$ as in (2). We say that $(Y_n)_{n\geq 0}$ linearizes $(X_n)_{n\geq 0}$ (or $\tilde{\mathcal{P}}$ linearizes $\mathcal{P}$) if the following two property holds: (i) if $Y_0 = (e,1)$ there exists a sequence of stopping times $(\tau_n)_{n\geq 0}$ with $\tau_0 = 0$ such that $(Y_{\tau_n})_{n\geq 0}$ is a realization of the random walk $(X_n)_{n\geq 0}$ with initial condition e, and (ii) these stopping times are a renewal process, that is the variables $\tau_{n+1} - \tau_n$ are iid with finite mean.*

To be precise, in the above definition, when saying that Y_{τ_n} is a convolution random walk, we of course identify Y_{τ_n} with its G coordinate and forget about the color, which in this case is constant equal to 1.

We remark also that because of a transitivity property, namely one can always translate the starting point to $e \in G$ by left multiplication, there is no loss of generality in supposing that the walks are started at $e \in G$. More details on this are given in Section 3.

Theorem 1. *Let G be a a group generated by a finite set S, with identity element e. Consider a random walk with kernel $\mathcal{P}$ as in (1) with finite support generating G. Then there exists $r \geq 1$ and a colored random walk on $G \times [r]$ with kernel $\tilde{\mathcal{P}}$ as in (2) which linearizes $\mathcal{P}$. Moreover, if $\mathcal{P}$ is reversible, then $\tilde{\mathcal{P}}$ can be chosen to be also reversible.*

Theorem 1 has been stated in a non-constructive manner for ease of notation. The proof of Theorem 1 in Section 2 will exhibit two simple linearization constructions which have a vector $p = (p_g)_{g \in G}$ as input and gives as output the integer r and the family of matrices $(\tilde{p}_g)_{g \in S}$. There is one construction in the general case and one construction which preserves reversibility. We refer to Remark 4 for the number of colors r needed in both constructions. We note also that the previous spectral linearization tricks reviewed in [13] did not preserve the Markov property and could not be used to prove Theorem 1.

There are possible extensions of the Markovian linearization trick. It is possible with the same proof techniques to linearize a colored random walk on $G \times [r]$ with finite range and obtain a colored nearest-neighbor colored random walk on $G \times [r']$ with $r' \geq r$. In another direction, it is possible to linearize random walks on G which do not have a finite range provided that we allow a countable set of colors, see Remark 2 below. Finally, in this paper we focus on groups only but our first linearization construction applies to random walks on monoids as well.

Application to the computation of drift and entropy. Initially, our goal was to use the linearization trick to try to compute explicit formulas for various invariant quantities of interest, mainly entropy and drift.

Elements of G can be written as products of elements in S, which will be called words. A word $a_1 \cdots a_k$ with the $a_i \in S$ such that $g = a_1 \cdots a_k$ is called a representative of the element $g \in G$. The length with respect to S of $g \in G$ is

$$|g| := \min\{k, g = a_1 \cdots a_k, a_i \in S\}. \tag{4}$$

Consider a convolution random walk $(X_n)_{n \geq 0}$ on a finitely generated group G, started at e. Assuming finite moment $\mathbb{E}[|X_1|] < \infty$, which is always satisfied if the walk is finitely supported, the sequence $(\mathbb{E}[|X_n|])_{n \geq 0}$ is easily shown to be sub-additive (see Section 3), which implies the existence of the limit

$$\gamma := \lim_{n \to \infty} \frac{\mathbb{E}|X_n|}{n}.$$

The scalar γ is called the drift, or rate of escape, of the random walk. The drift γ is in fact also an a.s. and L^1 limit, as can be shown using for instance Kingman's sub-additive theorem. Therefore

$$\gamma \coloneqq \lim_{n\to\infty} \frac{|X_n|}{n} \quad \text{a.s. and in } L^1.$$

Similarly, from the definition of the entropy of X_n as

$$H(X_n) \coloneqq -\sum_g \mathbb{P}[X_n = g] \log \mathbb{P}[X_n = g]$$

one can define

$$h \coloneqq \lim_{n\to\infty} \frac{H(X_n)}{n}.$$

as an L^1 limit or as an a.s. limit:

$$h = \lim_{n\to\infty} -\frac{\log \mathcal{P}^n(e, X_n)}{n} \quad \text{a.s. and in } L^1$$

where $\mathcal{P}^n$ the n-th power of the transition kernel of the random walk. h is called the Avez entropy, or entropy rate, or asymptotic entropy, of the random walk. In fact, this entropy can be interpreted as the drift for the Green pseudo-metric, see [1]. In this paper it will often simply be referred to as the entropy of the random walk. We refer the reader to Woess [15] for a general reference on the topic.

These notions can be extended to colored random walks: it is straightforward to extend the definition of $H(X_n)$, while for the drift we set $|x| \coloneqq |g|$ for all $x = (g, u) \in G \times [r]$.

Proposition 2. *Let $(X_n)_{n\geq 0}$ be a quasi-irreducible colored random walk on a group G. The following limits exist and do not depend on the starting color:*

$$\gamma = \lim_{n\to\infty} \frac{|X_n|}{n} \quad \textit{a.s. and in } L^1, \tag{5}$$

$$h = \lim_{n\to\infty} -\frac{\log \mathcal{P}^n((e,u), X_n)}{n} \quad \textit{a.s. and in } L^1. \tag{6}$$

for all starting color $u \in [r]$.

The proof is given to Section 3. The a.s. convergence of the entropy and drift combined with the law of large numbers yield the following corollary.

Corollary 3. *Let $(X_n)_{n\geq 0}$ be a finitely supported random walk on a group G and $(Y_n)_{n\geq 0}$ a colored random walk that linearizes $(X_n)_{n\geq 0}$ in the sense of Definition 3. The drift $\tilde{\gamma}$ and entropy $\tilde{h}$ of $(Y_n)_{n\geq 0}$ can be related to the drift γ and entropy h of $(X_n)_{n\geq 0}$ by:*

$$\gamma = \mathbb{E}[\tau_1]\tilde{\gamma}, \tag{7}$$

$$h = \mathbb{E}[\tau_1]\tilde{h}. \tag{8}$$

The expected time $\mathbb{E}[\tau_1]$ has a simple expression in the two linearization constructions given in Section 2, see Remark 1 and Remark 3.

In this paper, we have chosen to focus on drift and entropy but there are other numerical invariants associated to a random walk such as the spectral radius of $\mathcal{P}$ or the index of exponential decay of the return probability: $\lim_{n\to\infty} \mathcal{P}^{2n}(e,e)^{1/(2n)}$ (they coincide for reversible walks). Our linearization technique could also be used to study these last two quantities but this is less novel since these quantities can be read on the resolvent operator and previous linearization techniques allow to compute resolvent of operators of the form (1), see [13] for such linearization and [19, 6] for examples of computation of resolvent operators of the form (2).

Entropy and drift for plain groups. Most of the time the exact computation of the drift and entropy is not accessible. However for nearest-neighbor random walks on groups with an underlying tree structure like free groups [15, 18] and free products of finite groups [14, 21, 9, 7], there exist techniques that yield explicit formulas. This paper shows how to extend extend some of these techniques to the colored setting. Combined with the linearization trick, our results give explicit formulas for finitely supported random walks. By 'explicit', we mean that we express drift and entropy in terms of the unique solutions of some finite dimensional fixed point equations.

Following [14], we combine the case of free groups and free product of finite groups. These groups are called plain groups in [14, 12].

Definition 4. *A plain group is a free product of a finitely generated free group and a finite family of finite groups.*

Write $\mathbb{F}_d$ for the free group on d generators $a_1, \dots a_d$ and let $G_1, \dots, G_m$ be finite groups. Consider the plain group $G = \mathbb{F}_d * G_1 * \cdots * G_m$, with the set of generators

$$S := \bigcup_{i=1}^{d} \{a_i, a_i^{-1}\} \bigcup \left(\bigsqcup_{j=1}^{m} S_j \right).$$

In the above expression, for all $j = 1, \dots, m$, $S_j := G_j \setminus \{e\}$ is the set of elements of G_j distinct from the identity element, and $\bigsqcup$ denotes a disjoint union.

Introducing the notation from [14],

$$\forall g \in S, \quad \mathrm{Next}(g) := \begin{cases} S \setminus \{g^{-1}\} & \text{if } g \in \mathbb{F}_d \\ S \setminus S_i & \text{if } g \in S_i \end{cases}, \tag{9}$$

we see that every element $g \in G$ writes as a word $g = g_1 \cdots g_n$ with $n = |g|$ and $g_{i+1} \in \mathrm{Next}(g_i)$ for all $i = 1, \dots, n-1$.

If we set aside the trivial cases where G is isomorphic to $\mathbb{Z}$ or $\mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}$, G is a non-amenable group which implies any irreducible convolution random walk on G is transient (see for instance [25, Chapter 1]). Therefore if X_n is a convolution random walk on G, $|X_n| \to \infty$ and X_n converges in some sense to an infinite word X_∞ made of letters in S. The law p^∞ of X_∞ is called the harmonic measure and it provides much information on the asymptotic properties of X_n. It can in particular be used to compute its entropy and drift. Given $g \in S$, let $\mu_g := p^\infty[\xi_1 = g]$ be the probability under p^∞ that an infinite word starts with letter g. The most explicit formulas for nearest-neighbor random walks on plain groups express the entropy and drift in terms of the parameters p_g and μ_g.

In this paper we extend this technique to the colored setting. It had already been shown that such techniques could be adapted to nearest-neigbhor random walks on more general free products, see notably [23, 9, 10]. Here, the non-commutativity of $M_r(\mathbb{C})$ creates new difficulties. We consider a natural generalization of the harmonic measure p_u^∞, $u \in [r]$ being the starting color, and a matrix version of the parameters μ_g, which satisfy

$$p_u^\infty(\xi_1 = g) = \sum_{v \in [r]} \mu_g(u,v), \qquad \forall u \in [r].$$

The family of parameters $(\mu_g)_{g \in S}$ is uniquely characterized by a family of matrix relations (22), called the traffic equations, which arrives as consequences of the stationarity. The traffic equations can be solved quite explicitly for the free group. The precise definitions and properties of the harmonic measure are given in Section 4.

We start with a formula for the drift. In the sequel, $\mathbb{1}$ denotes the vector $\mathbb{R}^r$ with all coordinates equal to 1 and $\mathbb{1}_u$ the indicator at $u \in [r]$, that is the vector with zero coordinates except the u entry equal to 1. The usual scalar product is denoted by $\langle\ ,\ \rangle$.

Theorem 4. *Let $(Y_n)_{n \geq 0}$ be a nearest-neighbor quasi-irreducible colored random walk on $G \times [r]$ defined by a family of matrices $(p_g)_{g \in S}$. Let π be the unique invariant probability measure of the stochastic matrix $P = \sum_g p_g$. The drift of $(Y_n)_{n \geq 0}$ is given by*

$$\begin{aligned}\gamma &= \sum_{g \in S} \Big\langle \mathbb{1}, \pi p_g \Big(-\mu_{g^{-1}} + \sum_{h \in \mathrm{Next}(g)} \mu_h \Big) \Big\rangle. \\ &= \sum_{g \in S} \sum_{u,v,w \in [r]} \pi(u) p_g(u,v) \Big(-\mu_{g^{-1}}(v,w) + \sum_{h \in \mathrm{Next}(g)} \mu_h(v,w) \Big).\end{aligned}$$

Theorem 4 is the exact counterpart of the already known formulas for colorless walks (corresponding to $r = 1$).

For the entropy, we first provide an integral formula in Theorem 5 which is again the counterpart of a known formula in the colorless setting.

Theorem 5. *Let $(Y_n)_{n\geq 0}$ be a nearest-neighbor quasi-irreducible colored random walk on $G\times[r]$ defined by a family of matrices $(p_g)_{g\in S}$. Let π be the unique invariant probability measure of the stochastic matrix $P=\sum_g p_g$. The asymptotic entropy of $(Y_n)_{n\geq 0}$ is given by*

$$h=-\sum_{\substack{g\in G\\ u,v\in[r]}}\pi(u)p_g(u,v)\int\log\left(\frac{dg^{-1}p_u^\infty}{dp_v^\infty}\right)dp_v^\infty. \tag{10}$$

For nearest-neighbor convolution random walks on plain groups, it is then possible to derive an explicit expression from this integral formula. For colored random walks, adapting the computation naturally leads to the problem of determining the limit of infinite products of random matrices, which in turn leads to a formula of the entropy in terms of the parameters $(\mu_g)_{g\in S}$ and the law of a family indexed by $[r]$ of random probability measures on $\mathbb{R}^r$ whose law is uniquely characterized by some invariance equation.

To state our result more precisely, we first introduce the hitting probabilities defined as follows. Given $g\in G, u\in[r]$, let $\tau_{(g,u)}:=\inf\{n: X_n=(g,u)\}$ be the hitting time of a pair (g,u), so that $\tau_g:=\min_u\tau_{(g,u)}$ is the hitting time of g. For $g\in S$, set

$$q_g(u,v):=\mathbb{P}[\tau_g<\infty\text{ and }\tau_g=\tau_{(g,v)}]. \tag{11}$$

As we will check, the matrices q_g and μ_g satisfy a simple relation (28) and the hitting probabilities $(q_g)_{g\in S}$ are characterized by a quadratic equation (27) which can be solved quite explicitly for the free group. In the statement below, $\mathcal{P}_+$ denotes the open simplex of probability measures $[r]$:

$$\mathcal{P}^+:=\left\{x\in\mathbb{R}^r:\sum_{i\in[r]}x_i=1\text{ and }x_i>0,\ \forall i\in[r]\right\}. \tag{12}$$

Theorem 6. *Suppose the matrices $(q_g)_{g\in S}$ satisfy (43) defined in Section 3. Then the entropy of the colored random walk is given by*

$$\begin{aligned}h=-\sum_{g\in S}\sum_{u,v,w\in[r]}\pi(u)p_g(u,v)\Bigg[&\mu_{g^{-1}}(v,w)\int\log\left(\frac{\langle\mathbb{1}_v\,,\,z\rangle}{\langle q_{g^{-1}}(u,\cdot)\,,\,z\rangle}\right)d\nu_w(z)\\&+\sum_{h\in\mathrm{Next}(g)}\mu_h(v,w)\int\log\left(\frac{\langle q_g(v,\cdot)\,,\,q_hz\rangle}{\langle\mathbb{1}_u\,,\,q_hz\rangle}\right)d\nu_w(z)\\&+\sum_{h\in S:gh\in S}\mu_h(v,w)\int\log\left(\frac{\langle q_{gh}(v,\cdot)\,,\,z\rangle}{\langle q_h(u,\cdot)\,,\,z\rangle}\right)d\nu_w(z)\Bigg],\end{aligned} \tag{13}$$

where $(\nu_u)_{u\in[r]}$ is the unique family of probability measures on $\mathcal{P}^+$ satisfying

$$\int f(z)\,d\nu_u(z)=\sum_{\substack{g\in S\\ v\in[r]}}\int f(q_gz)\mu_g(u,v)\,d\nu_v(z),\quad\forall u\in[r],$$

for all bounded measurable functions f on $\mathcal{P}^+$.

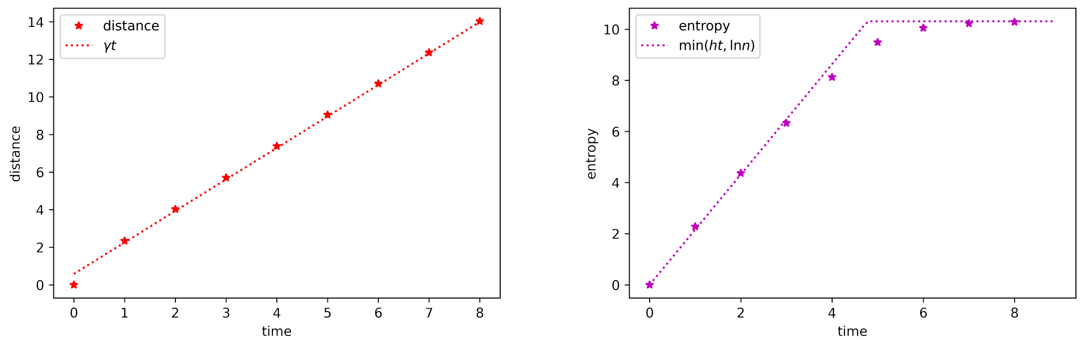

Figure 1: Numerical simuation of a finite range random walk $(X_t)_{t\geq 0}$ on $\mathbb{F}_2 = \langle a, b\rangle$ with 12 generators of length $1, 2$ and 3, and with random transition probabilities. The linearization procedure gives $r = 16$. On the left, an average of $|X_t|$ over $n = 20000$ trajectories compared to the line γt. On the right, an average of $H(X_t)$ over n trajectories compared to the line ht. For computational efficiency, the entropy $H(X_t)$ is approximated with the entropy of a random walk on $[n]$ where $\lambda(g), g \in \{a, b\}$, are replaced by an independent uniform permutation matrix of size n (in other words, we consider a random action of $\mathbb{F}_2$ on a set of size n). This approximation is valid as long as $ht \leq \ln n$, see [2].

The technical condition (43) is described in Section 3. Let us simply note that this condition is automatically satisfied for colored random walks arising as a linearization of a finite range random walk on G (this is the content of Proposition 17 below). Applying Theorem 6 to a nearest-neighbor walk for which $r = 1, \pi = 1, \nu = \delta_1$, we get back formula (22) in [14]:

$$h = -\sum_{g\in S} p_g \left(\mu_{g^{-1}} \log \frac{1}{q_{g^{-1}}} + \sum_{h\in \mathrm{Next}(g)} \mu_h \log q_g + \sum_{h\in S:\ gh\in S} \mu_h \log \frac{q_{gh}}{q_h} \right).$$

As alluded to above, in Theorem 6, the measure ν_g on $\mathcal{P}_+$ will arise as the convergence in direction toward a random rank one projector of the product of matrices $q_{g_1} q_{g_2} \cdots q_{g_n}$ where $g_1, g_2, \ldots$ in S are the successive letters of X_∞ with law p^∞ conditioned to start with $g_1 = g$.

For finite range random walks on plain groups, fine qualitative results on drift and entropy were already available such as [16, 17] and even on general hyperbolic groups [11, 8] and references therein. As pointed in [18, 8], the computations for drift and entropy for nearest-neighbor random walk on the free group have been known for around 50 years but it was however unexpected that these explicit formulas could be extended to finite range random walks. To our knowledge, Theorem 4 and Theorem 6 provide the first formulas for finite range random walks on plain groups, see Figure 1 for a numerical illustration.

Organization of the paper. In Section 2 we propose a general construction of linearized colored walks, with some variations, that lead to Theorem 1. Section 3 is devoted to the proof of Proposition 2. In Section 4 we define the harmonic measure for a random walk on a plain group. Mimicking results for standard random walks, we prove the Markovian structure of the harmonic measure, which is the main ingredient for computing drift and entropy. The computations leading to Theorem 4 and Theorem 5 and additionnal formulas for the entropy are given in section 5. Finally Section 6 provides a modest example illustrating the use of the linearization trick.

Acknowledgments. Part of this work was performed during a visit of both authors to Kyoto University. Both authors gratefully acknowledge the support of JSPS and Kyoto University. CB was supported by the research grant ANR-16-CE40-0024-01.

2 Linearization: proof of Theorem 1

Let $(X_n)_{n\geq 0}$ be a finitely supported random walk on G. In this section we construct a nearest-neighbor colored random walk $(Y_n)_{n\geq 0}$ from $(X_n)_{n\geq 0}$ in such a way the entropy and drift can easily be related. What seems the most natural is to directly look for Y_n as the walk that follows the same trajectory as X_n but decomposes a transition in several steps in order to be nearest-neighbor. What follows formalizes this idea and adds two possible improvements. The first improvement allows to reduce the dimension of the matrices, the second is a linearization that preserves reversibility.

From now on, let $a_1, \dots, a_d$ be the elements of the generating set S, and $a_0 := e$. Let K denote the support of the probability measure p and ℓ_K be the maximal length of an element of K. For each $g \in K$ choose a representative $s(g)$, which can always be assumed to be of length $|g|$. We will generally identify g with its representative and write simply $g = a_{i_1} \cdots a_{i_k}$ to refer to the fact that $s(g) = a_{i_1} \cdots a_{i_k}$, for some $i_1, \dots, i_k \in [d]$. Similarly, set $p_{i_1 \cdots i_k} := p_g$, with in particular $p_0 := p_e$ and $p_i := p_{a_i}$, whether $a_i \in K$ or not. Since we look for a nearest-neighbor colored random walk, the colored random walk $(Y_n)_{n\geq 0}$ is defined by the $r \times r$ matrices $\tilde{p}_i, i \in [0:d]$.

Initial construction. As mentioned previously, one can first construct Y_n as the walk which, when X_n is multiplied by $g = a_{i_1} \cdots A_{i_k} \in K$, visits each a_i consecutively. Colors can be used to impose Y_n to follow the trajectory of the initial random walk X_n. For instance, consider a neutral color, written 1 and given $g = a_{i_1} \cdots a_{i_k} \in K$, consider $k-1$ different colors $u_1(g), \dots, u_{k-1}(g)$ and set $\tilde{p}_{i_1}(1, u_1(g)) := p_g, \tilde{p}_{i_2}(u_1(g), u_2(g)) := 1, \dots, \tilde{p}_{i_k}(u_k(g), 1) = 1$. Here the color 1 is to be interpreted as a neutral color: supposing the colored walk Y_n is at

$(e,1)$, it goes to a_{i_1} with probability $p_{i_1\cdots i_k}$, and is then forced to go through $a_{i_2}\cdots a_{i_k}$ until it is back at color 1. Except for the color 1, all the colors serve only for the intermediary steps. If one considers now another element $h \in K, h = a_{i'_1}\cdots a_{i'_{k'}}$, one can take $k'-1$ new colors to distinguish between the intermediary steps which lead to g and those which lead to h. Repeating the process with each generator in the support of the walk, one can construct a colored random walk $(Y_n)_{n\geq 0}$ such that $Y_{\tau_n} \overset{(d)}{=} X_n$ for all n, where τ_n are the successive hitting times of the color 1.

2.1 Linearization in the non-reversible case

We now describe more precisely a construction which reduces the number of colors needed in the initial construction by noticing that elements with common prefixes can use the same colors when visiting these prefixes.

Recall that a representative of minimal length has been fixed for each element of K. These representatives are called the representatives of K. Given $g,h \in K, g = a_{i_1}\cdots a_{i_k}, h = a_{j_1}\cdots a_{j_{k'}}$, the prefix of g and h is defined as $g \wedge h = a_{i_1}\cdots a_{i_m}$ where $m := \max\{n \geq 0 : i_1 = j_1, \dots, i_n = j_n\}$. h is a strict prefix of g if $g \wedge h = h$ and $|g| > |h|$. For $k \in [\ell_K]$ and $i_1 \dots, i_k \in [d]$, let

$$[i_1\cdots i_k] := \{g \in K, g \wedge a_{i_1}\cdots a_{i_k} = a_{i_1}\cdots a_{i_k}, |g| \geq k+1\}$$

and

$$p_{[i_1\cdots i_k]} := \sum_{g\in[i_1\cdots i_k]} p_g$$

be the cumulative mass under p of all words in K for which $a_{i_1}\cdots a_{i_k}$ is a strict prefix. Then for $k \in [\ell_K]$, we set

$$q_{i_1\cdots i_k} = \frac{p_{[i_1\cdots i_k]}}{p_{[i_1\cdots i_{k-1}]}}$$

where the denumerator is to be taken as equal to one in the case $k = 1$.

Here we will make use of the operator point of view in order to write all the matrices defining $(Y_n)_{n\geq 0}$ at once. As $M_r(\mathbb{C}) \otimes \mathcal{A}$ is isomorphic to $M_r(\mathcal{A})$ for any unital algebra $\mathcal{A}$, the transition kernel $\tilde{P}$ can be written as one matrix whose coefficients are operators on $\ell^2(G)$. In this case, the matrix has coefficients which are linear combinations of the multiplication operators. The matrices $p_i, i \in [0,d]$ can easily be deduced: to obtain p_i it suffices to replace $\lambda(a_i)$ by 1 and the $\lambda(a_j), j \neq i$ by 0.

For all $k \in [\ell_K]$, let

$$r_k := \text{Card } \{(i_1, \ \dots \ , i_k) \in S^k, [i_1\cdots i_k] \neq \varnothing\}$$

be the number of strict prefixes of elements of K that have length k. Then, we define the following matrices: $C(k)$ is the $r_{k-1} \times 1$ column matrix

$$C(k) := \left(\sum_{j \in [d]} \frac{p_{i_1 \cdots i_{k-1} j}}{p_{[i_1 \cdots i_{k-1}]}} \lambda(a_j) \right)_{i_1 \cdots i_{k-1}},$$

which is indexed by strict prefixes of K of length $k-1$. Given a representative $i_1 \cdots i_{k-1}$, define the row matrix

$$L(i_1, \ldots, i_{k-1}) := (q_{i_1 \cdots i_{k-1} j} \, \lambda(a_j))_j$$

indexed by all j such that $i_1 \cdots i_{k-1} j$ is the strict prefix of a representative of K. Then use these row matrices to form the $r_{k-1} \times r_k$ diagonal block matrix $D(k)$, whose $i_1 \cdots i_{k-1}$ diagonal entry is $L(i_1, \cdots, i_{k-1})$:

$$D(k) = \begin{pmatrix} \ddots & & \\ & L(i_1, \ldots, i_{k-1}) & \\ & & \ddots \end{pmatrix}.$$

Finally combine all the preceding matrices to construct:

$$\tilde{\mathcal{P}} := \begin{pmatrix} \sum_{i \in [0:d]} p_i \lambda(a_i) & D(1) & 0 & & & 0 \\ C(2) & 0 & D(2) & & & \\ C(3) & 0 & 0 & D(3) & & \\ \vdots & & & & \ddots & \\ C(\ell_K - 1) & & & & & D(\ell_K - 1) \\ C(\ell_K) & & & & & 0 \end{pmatrix} \tag{14}$$

Example 1. *The construction will certainly be clearer on a concrete example. Consider a random walk on the group G given by the presentation $\langle a, b \mid ab = ba \rangle$, which is in fact isomorphic to $\mathbb{Z}^2$. Suppose the random walk is supported by all words in a and b of length less than 3, that is*

$$K = \{e, a, b, a^2, ab, b^2, a^3, a^2 b, ab^2, b^3\}.$$

To avoid large matrices, we forget about the inverses a^{-1}, b^{-1} but they should not be discarded in a general case.

Because of the relation $ab = ba$, there can be several ways to write an element of G as words in a and b, for instance $bab = ab^2$. Therefore we fix a representative for each element of the support. In the present case, group elements can be written uniquely as $a^k b^l$ so it is natural to choice these words as representatives.

Applying the preceding construction, one eventually obtains the following operator matrix:

$$\begin{pmatrix} p_e\lambda(e)+p_a\lambda(a)+p_b\lambda(b) & q_a\lambda(a) & q_b\lambda(b) & 0 & 0 & 0 \\ \frac{p_{a^2}}{p_{[a]}}\lambda(a)+\frac{p_{ab}}{p_{[a]}}\lambda(b) & 0 & 0 & q_{a^2}\lambda(a) & q_{ab}\lambda(b) & 0 \\ \frac{p_{b^2}}{p_{[b]}}\lambda(b) & 0 & 0 & 0 & 0 & q_{b^2}\lambda(b) \\ \frac{p_{a^3}}{p_{[a^2]}}\lambda(a)+\frac{p_{a^2b}}{p_{[a^2]}}\lambda(b) & 0 & 0 & 0 & 0 & 0 \\ \frac{p_{ab^2}}{p_{[ab]}}\lambda(b) & 0 & 0 & 0 & 0 & 0 \\ \frac{p_{b^3}}{p_{[b^2]}}\lambda(b) & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

where

$$q_a = p_{[a]} = p_{a^2} + p_{ab} + p_{a^3} + p_{a^2b} + p_{ab^2}, \quad q_b = p_{[b]} = p_{b^2} + p_{b^3}, \quad q_{a^2} = p_{[a^2]}/p_{[a]} = (p_{a^3} + p_{a^2b})/p_{[a]}$$

$$q_{ab} = p_{[ab]}/p_{[a]} = p_{ab^2}/p_{[a]}, \quad q_{b^2} = p_{[b^2]}/p_{[b]} = p_{b^3}/p_{[b]}.$$

Proof of Theorem 1 (non-reversible case). On row $i_1\cdots i_k$, the sum of entries of the matrix $P = \sum \tilde{p}_i$ is

$$\sum_{j\in[d]} \frac{p_{i_1\cdots i_{k-1}j}}{p_{[i_1\cdots i_{k-1}]}} + \sum_{j\in[d]} q_{i_1\cdots i_{k-1}j} = \frac{1}{p_{[i_1\cdots i_{k-1}]}} \sum_{j\in[d]} (p_{i_1\cdots i_{k-1}j} + p_{[i_1\cdots i_{k-1}j]}) = 1.$$

Thus P is stochastic and $\tilde{\mathcal{P}}$ defines indeed a colored random walk $(Y_n)_{n\geq 0}$.

Suppose now $Y_n = (g_n, u_n)$ is started at color 1. Then define $(\tau_n)_{n\geq 0}$ as the successive return times at the first color: $\tau_0 := 0$ and for $n \geq 1$

$$\tau_n := \inf\{m > \tau_{n-1}, u_m = 1\}$$

By the Markov property, the random variables $\tau_n - \tau_{n-1}$ are iid with the same law as τ_1.

For all $g = i_1\cdots i_k$ in K, the probability that $Y_{\tau_1} = g$ is

$$\begin{aligned}\mathbb{P}\left[Y_{\tau_1} = g\right] &= q_{i_1}\, q_{i_1 i_2} \cdots\, q_{i_1\cdots i_{k-1}}\, \frac{p_{i_1\cdots i_k}}{p_{[i_1\cdots i_{k-1}]}} \\ &= p_{[i_1]} \frac{p_{[i_1 i_2]}}{p_{[i_1]}} \cdots \frac{p_{[i_1\cdots i_{k-1}]}}{p_{[i_1\cdots i_{k-2}]}} \frac{p_{i_1\cdots i_k}}{p_{[i_1\cdots i_{k-1}]}} \\ &= p_{i_1\cdots i_k} = p_g.\end{aligned}$$

By Markov property and the fact that the increments are iid one easily deduce that $Y_{\tau_n} \overset{(d)}{=} X_n$ for all $n \geq 0$. Finally, the irreducibility of X_n implies the quasi-irreducibility of Y_n. □

Remark 1. The expectation of the hitting time τ_1 is very simple to compute: Y_1 can go through each $g \in K$ with probability p_g, in which case it needs $|g|$ steps to get back to the color 1. Hence the expectation of τ_1 is just the average length of elements in K:

$$\mathbb{E}\left[\tau_1\right] = p_e + \sum_{g\in K} p_g |g| = p_e + \mathbb{E}[|X_1|].$$

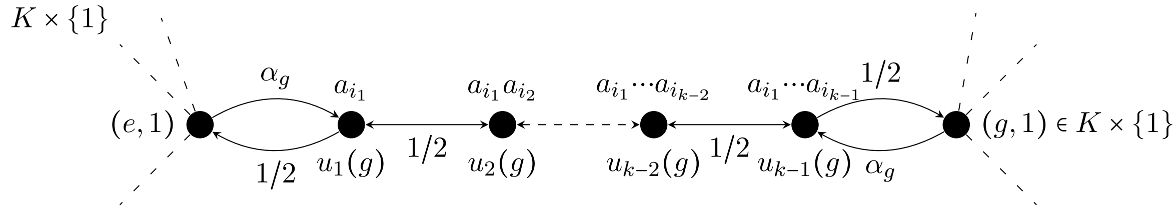

Figure 2: The linearizing reversible random walk

Remark 2. If $\mathcal{P}$ is not of finite range, then the construction produces a countable set of colors. The construction could nevertheless be useful if the expected time $\mathbb{E}[\tau_1] = p_e + \mathbb{E}[|X_1|]$ is finite.

2.2 Linearization in the reversible case

In the previous constructions, it is impossible for the colored random walk to backtrack and hence to be reversible.

To correct this, we propose the following variation of the initial construction. We assume that for all $g \in K$, the representative of $g = a_{i_1} \cdots a_{i_n}$ is chosen such that the representative of g^{-1} is $a_{i_n}^{-1} \cdots a_{i_1}^{-1}$. We start with the neutral color 1 and, for each pair (g, g^{-1}) with $|g| = |g^{-1}| \geq 2$, we add $|g| - 1$ new colors $u_1(g), \ldots, u_{|g|-1}(g)$ and we set $u_k(g^{-1}) = u_{|g|-k}(g)$ for all $1 \leq k \leq |g| - 1$. Suppose g is written as $g = a_{i_1} \cdots a_{i_k}$. For all $h \in G$, the transition probability to go from $(h, 1)$ to $(ha_{i_1}, u_1(g))$ is set to some value α_g to be determined, such that $\alpha_g = \alpha_{g^{-1}}$ and $\sum_{g \in S} \alpha_g = 1$. All other transition probabilities on the segment joining $(x, 1)$ to $(xg, 1)$ are set to $1/2$, see Figure 2.

In matrix form, the construction is as follows. Given $\alpha \in \mathbb{R}_+$ and $g = a_{i_1} \cdots a_{i_k}, k \geq 3$, consider the following operator-valued matrix of size $k \times k$

$$A_g(\alpha) := \begin{pmatrix} 0 & \alpha\lambda(a_{i_1}) & 0 & \cdots & 0 & \alpha\lambda(a_{i_k}^{-1}) \\ 1/2\,\lambda(a_{i_1}^{-1}) & 0 & 1/2\,\lambda(a_{i_2}) & 0 & \cdots & 0 \\ 0 & 1/2\,\lambda(a_{i_2}^{-1}) & 0 & \ddots & & \vdots \\ \vdots & 0 & \ddots & 0 & & 0 \\ 0 & \vdots & & 1/2\,\lambda(a_{i_{k-2}}^{-1}) & 0 & 1/2\,\lambda(a_{i_{k-1}}) \\ 1/2\,\lambda(a_{i_k}) & 0 & \cdots & 0 & 1/2\,\lambda(a_{i_{k-1}}^{-1}) & 0 \end{pmatrix}.$$

If $k = 2$, let

$$A_{a_{i_1}a_{i_2}}(\alpha) := \begin{pmatrix} 0 & \alpha\lambda(a_{i_1}) + \alpha\lambda(a_{i_2}^{-1}) \\ 1/2\,\lambda(a_{i_2}) + 1/2\,\lambda(a_{i_1}^{-1}) & 0 \end{pmatrix}.$$

If $k = 1$, let $A_{a_i}(\alpha) = (\alpha\lambda(a_i) + \alpha\lambda(a_i^{-1}))$ and finally for $k = 0$ set $A_e(\alpha) := (\alpha\lambda(e))$. In the case $k \geq 2$, one can write

$$A_g(\alpha) = \begin{pmatrix} 0 & L_{i_1,i_k}(\alpha) \\ C_{i_1,i_k} & \tilde{A}_{i_2\cdots i_{k-1}} \end{pmatrix},$$

where $L_{i_1,i_k}(\alpha), C_{i_1,i_k}$ are respectively a $1 \times (k-1)$ row matrix and a $(k-1) \times 1$ column matrix, while $\tilde{A}_{i_2\cdots i_{k-1}}$ is a tri-diagonal $(k-1) \times (k-1)$ matrix depending only $i_2, \ldots, i_{k-1}$, with zero diagonal. Now given a square matrix M of matrix $m \times m$, use $A_g(\alpha)$ to form the matrix of size $m + k - 1$

$$M \cup A_g(\alpha) := \begin{pmatrix} \begin{matrix} & M & \end{matrix} & \begin{matrix} L_{i_1,i_k}(\alpha) \\ 0 \end{matrix} \\ \begin{matrix} C_{i_1,i_k} & 0 \end{matrix} & \tilde{A}_{i_2,\ldots,i_{k-1}} \end{pmatrix},$$

where 0 is the null matrix of the appropriate size. If $k \in \{0,1\}$, simply set $M \cup A_g(\alpha) := M + \mathrm{diag}(A_g(\alpha), 0, \ldots, 0)$.

The transition kernel eventually considered is

$$\tilde{\mathcal{P}} := \bigcup_{g \in K} A_g(\alpha_g). \tag{15}$$

Thanks to the constraint $\sum_{g \in G} \alpha_g = 1$, $\tilde{\mathcal{P}}$ is indeed a Markov transition kernel. However, the colored walk can now make steps that would not have been possible in the initial construction. For instance, it can go from $(e,1)$ to some (g,u) and then come back to $(e,1)$, even if $p_e = 0$ for the initial walk. Therefore, the stopping times cannot be taken directly as the return times at color 1. Instead define, if $Y_n = (g_n, u_n)$ has transition kernel $\tilde{\mathcal{P}}$,

$$\tau_n := \begin{cases} \inf\{m \geq \tau_{n-1} : u_m = 1, g_m \neq g_{\tau_{n-1}}\} & \text{if } g_{\tau_{n-1}+1} \neq g_{\tau_{n-1}} \\ \tau_{n-1} + 1 & \text{if } g_{\tau_{n-1}+1} = g_{\tau_{n-1}} \end{cases}.$$

Theorem 1 in the reversible case is now contained in the following lemma.

Lemma 7. *Suppose X_n is a finitely supported random walk on G defined by a probability vector p and started at e. Setting for each $g \in K$*

$$\alpha_g := \begin{cases} \frac{(1-p_e)|g|p_g}{\sum_{h \in K} |h| p_h} & \textit{if } g \neq e \\ p_e & \textit{if } g = e \end{cases} \tag{16}$$

the preceding construction yields an operator $\tilde{\mathcal{P}}$ which defines a reversible colored random walk $(Y_n)_{n \geq 0}$ satisfying $Y_{\tau_n} \overset{(d)}{=} X_n$ for all $n \geq 0$ if started at $(e,1)$.

Proof. As already pointed, the construction define a colored random walk as long as $\sum_{g \in S} \alpha_g = 1$, which is true for α_g defined as in (16). Furthermore such a colored walk is reversible: for all $(g,u) \in G \times [r]$, letting

$$\mu(g,u) = \begin{cases} 1 & \text{if } u = 1 \\ 2\,\alpha_h & \text{if } u = u_h(k) \text{ for some } k < |h|. \end{cases}$$

defines a reversible measure for the random walk, as it can be checked directly.

Consider now the colored walk $(Y_n)_{n\geq 0}$ defined by (16), started at $Y_0 = (e,1)$. Suppose first that $p_e = 0$, so that τ_1 is the hitting time of the set $(K \setminus \{e\}) \times \{1\}$.

At step 1 the colored walk necessarily enters the segment joining $(e,1)$ to $(h,1)$ for some $h \in K$. If the walk escapes this segment at h, then $Y_{\tau_1} = h$. Otherwise, it necessarily goes back to the starting state $(e,1)$. Now on each of these segments, the colored walk is simply performing a simple random walk so the escape probabilities are given by the standard gambler's ruin problem. Namely the simple random walk on $[0:n]$ reaches n before it gets to 0 with probability k/n when started at $k \in [0:n]$. Therefore by Markov's property

$$\mathbb{P}_{(e,1)}\left[Y_{\tau_1} = g\right] = \alpha_g/|g| + \sum_{h\in K} \alpha_h (1 - 1/|h|)\mathbb{P}_{(e,1)}\left[Y_{\tau_1} = g\right].$$

We deduce that

$$\mathbb{P}_{(e,1)}\left[Y_{\tau_1} = g\right] = \frac{\alpha_g/|g|}{\sum_{h\in K} \alpha_h/|h|} = p_g.$$

For the general case $p_e \neq 0$, $Y_{\tau_1} = (e,1)$ if and only if $Y_1 = (e,1)$ which occurs with probability $\alpha_e = p_e$. For $g \neq 0$, consider the random walk conditionned to move at every step, written $(Y'_n)_{n\geq 0}$. If $q(x,y)$ is the transition probability for Y_n between two states x and y, then the transition probability for Y'_n is 0 if $y = x$ and $q(x,y)/(1-q(x,x))$ otherwise. The previous argument apply to this walk, so

$$\mathbb{P}_{(e,1)}\left[Y'_{\tau'_1} = g\right] = \frac{\alpha_g/((1-p_e)|g|)}{\sum_{h\in K\setminus\{e\}} \alpha_h/((1-p_e)|h|)} = \frac{p_g}{1-p_e},$$

with τ'_1 is the obvious extension of τ_1 to Y'_n. Coming back to Y_{τ_1}: to reach $g \neq e$ it it is necessary that $Y_1 \neq e$, which occurs with probability $(1-p_e)$. Conditional on that event it no longer matters whether the random walk comes back at e and possibly stays there, so one can reason with Y'_n instead. Hence

$$\mathbb{P}_{(e,1)}\left[Y_{\tau_1} = g\right] = (1-p_e)\mathbb{P}_{(e,1)}\left[Y'_{\tau'_1} = g\right] = p_g.$$

The conclusion follows. □

Remark 3. The expected time $\mathbb{E}[\tau_1]$, although not as simple as in the non-reversible case, can nonetheless be computed quite easily using for instance the electric network paradigm. Using [20, Prop 2.20], we found in the case $p_e = 0$

$$\mathbb{E}[\tau_1] = \sum_{g\in K} p_g |g|^2 = \mathbb{E}|X_1|^2. \tag{17}$$

Remark 4. In the initial construction, we have that the total number of colors is

$$r = 1 + \sum_{g\in K}(|g| - 1).$$

In the reversible construction, there is a factor $1/2$ on front of the sum (because we use the same colors for g and g^{-1}). In the improved construction of Subsection 2.1, this is an upper bound, the actual value is $r = 1 + \mathrm{Card}(K') - \mathrm{Card}(K)$ where K' is the set of words which are a suffix of some element in K (in the chosen representatives). In a concrete application, it is often possible to design linearization procedures which are more economical in terms of number of colors, see Section 6 for an example.

3 Entropy and drift for colored random walks

In this section we extend the notions of entropy and drift to colored random walks, proving Proposition 2.

Transitivity properties. If $(X_n)_{n\geq 0}$ is a Markov chain with transition kernel Q on a discrete state space V, (X_n) is said to be transitive if for all $x, y \in V$ there exists a bijection $\varphi : V \to V$ such that $\phi(x) = y$ and for all $z \in V$, $Q(x,z) = Q(y, \phi(z))$. Transitivity has the consequence that the Markov chain can be translated in order to start on a specific state. Sometimes a weaker notion of transitivity is satisfied: there exists a finite partion of the state space such that a bijections as above exist only between states that are in the same set of the partition. In this case the Markov chain is said to be quasi-transitive.

Transitivity is an essential property of convolution random walks on groups, bijections being simply given by the left multiplications. It is the property that makes the sequences $(|X_n|)_{n\geq 0}$ and $(H(X_n))_{n\geq 0}$ sub-additive and hence allows to define drift and entropy.

On the other hand, for colored random walks it is possible that no bijection as above exists between pairs (g,u) and $(h,v) \in G \times [r]$, but there is one between (g,u) and (h,u) for all $g, h \in G, u \in [r]$, given by the left multiplication by hg^{-1}. Hence colored random walks are only quasi-transitive. For this reason it is slightly less straightforward to define entropy and drift for colored random walks than it is for convolution walks.

Let $(X_n)_{n\geq 0}$ be a colored random walk defined by a family of matrices $(p_g)_{g\in G}$. Thanks to quasi-transitivity it can be without loss of generality to be started at $e \in G$. Write $\mathbb{P}_u = \mathbb{P}_{(e,u)}$ for the law of this colored chain when started at (e,u).

The definitions of entropy and drift can naturally be extended to colored random walks: recall that if $X_n = (g_n, u_n) \in G \times [r]$, $|X_n| := |g_n|$ and set

$$H_u(X_n) := -\sum_{g,v} \mathbb{P}_u\left[X_n = (g,v)\right] \log \mathbb{P}_u\left[X_n = (g,v)\right].$$

Let π be the unique invariant probability measure of the stochastic matrix P. Let $\mathbb{P}_\pi$ denote the law of the colored Markov chain started at e with the starting color being chosen according to π, that is

$$\mathbb{P}_\pi = \sum_{u\in[r]} \pi(u)\mathbb{P}_u.$$

Then $H_\pi(X_n) = \sum_{u\in[r]} \pi(u)H_u(X_n)$ forms a sub-additive sequence. Indeed, for $k \leq n$, the Markov property yields

$$\begin{aligned} H_u(X_n \mid X_k) &= \sum_{g,v} \mathbb{P}_{(e,u)}(X_k = (g,v))H_v(X_{n-k}) \\ &= \sum_v P^k(u,v)H_v(X_{n-k}). \end{aligned}$$

If we multiply by $\pi(u)$ and sum over $u \in [r]$, this proves that

$$H_\pi(X_n \mid X_k) = H_\pi(X_{n-k}) \tag{18}$$

thanks to the invariance of π. Therefore one can bound

$$H_\pi(X_{n+m}) \leq H_\pi(X_{n+m}, X_m) = H_\pi(X_m) + H_\pi(X_{n+m}|X_m) = H_\pi(X_m) + H_\pi(X_n).$$

and prove the existence of the limit

$$h_\pi := \lim_{n\to\infty} \frac{H_\pi(X_n)}{n}.$$

The previous computation is very simple but only give limits in expectation. By applying Kingman's subadditive theorem one can obtain L^1 as well as a.s. limits. Because the color set $[r]$ is finite, one can also expect that the limits do not depend on the starting color in the case P is irreducible. This is exactly the content of Proposition 2. To prove this proposition, the following lemma will be needed:

Lemma 8 ([20, Lemma 14.11])**.** *Let $(X_n)_n$ be a Markov chain on a state space V and let $f : V^{\mathbb{N}} \to \mathbb{R}$ be a Borel function. If the law of $f = f(X_0, X_1, \dots)$ does not depend on X_0 and $f(X_1, X_2, \dots) \leq f(X_0, X_1, \dots)$ a.s., then f is a.s. constant.*

Proof of Proposition 2. We make use of the following explicit construction of the probability space. Let $\Omega := (G \times [r])^{\mathbb{N}}$, let $\mathcal{F}$ be the product sigma-algebra and $\theta : (\omega_n)_{n\geq 0} \mapsto (\omega_{n+1})_{n\geq 0}$ the shift operator. The coordinates of $\omega_i \in G \times [r]$ are written here $\omega_i = (g_i, u_i)$.

Consider the measure $\mathbb{P}_{g,u}$ on Ω defined on cylinders by

$$\mathbb{P}_{g,u}[\omega_0 = (g_0,u_0), \omega_1 = (g_1,u_1), \dots, \omega_n = (g_n,u_n)] = \mathbb{1}_{g_0=g, u_0=u}\; p_{g_1}(u,u_1)\cdots p_{g_n}(u_{n-1},u_n).$$

Then $X_n := (g_0 g_1 \cdots g_n, u_n)$ is a realization of the colored random walk defined by p and $\mathbb{P}_{g,u}$ is indeed the law of the random walk started at (g,u) where ω_n is the pair generator-color chosen by the walk at time n.

Thanks to the invariance of π with respect to P, the measure

$$\mathbb{P}_{\pi p} := \sum_{g,u} \pi(u) p_g(u,v) \mathbb{P}_{g,v}$$

is invariant by the shift θ. Let $\mathbb{P}^n_{g,u}$ denote the law of X_n under $\mathbb{P}_{g,u}$ and consider the function f_n on Ω defined by

$$f_n = -\log \mathbb{P}^n_{X_0} [X_n] .$$

By the Markov property, $\mathbb{P}^{m+n}_{X_0}[X_{m+n}] \geq \mathbb{P}^m_{X_0}[X_m]\mathbb{P}^n_{X_m}[X_{n+m}]$, so f_n is sub-additive in the sense that $f_{n+m} \leq f_m + f_n \circ \theta^m$. Because the walk is finitely supported, f_1 is integrable so Kingman's sub-additive theorem ensures that (f_n/n) converges both a.s. and in L^1 to a function $f \in L^1(\mathbb{P}_{\pi p})$, invariant under θ and such that $\int f \, d\mathbb{P}_{\pi p} = \lim_n \int f_n/n \, d\mathbb{P}_{\pi p} = \inf_n \int f_n/n \, d\mathbb{P}_{\pi p}$.

The last step is to use the lemma to prove that f is constant. By abuse of notation let us write $f = f(X_0, X_1, \ldots)$. Since $\mathbb{P}^n_{X_0}[X_n] \geq \mathbb{P}_{X_0}[X_1]\mathbb{P}^{n-1}_{X_1}[X_n]$, by taking logarithms and dividing by n, we get that a.s. $f(X_1, X_2, \ldots) \leq f(X_0, X_1, \ldots)$.

It then remains to see that the law of f does not depend on X_0. Thanks to quasi-transitivity, the law of f can only depend on the starting color. On the other hand the matrix P is irreducible so for all starting color u there exists a.s. a random integer m such that X_m has color v. Suppose m is the first time color v is visited. Then, by the strong Markov property, for all $n \geq 0$

$$\begin{aligned} \mathbb{P}^{n+m}_{e,u}[X_{n+m}] &\geq \mathbb{P}^m_{e,u}[X_m]\mathbb{P}^n_{g_m,v}[X_{n+m}] \\ &= \mathbb{P}^m_{e,u}[X_m]\mathbb{P}^n_{e,v}[\tilde{X}_n], \end{aligned}$$

with $\tilde{X}_n = \omega_{m+1} \cdots \omega_n$. Now take logarithms, divide by n and take $n \to \infty$. The left-hand side converges a.s. to $f((e,u), \ldots)$ while the right-hand side converges to $f((e,v), \ldots)$. Thus a.s. we have $f((e,u), \ldots) \geq f((e,v), \ldots)$. By symmetry the converse inequality also holds true so there is in fact equality, which proves f is constant. Hence the entropy is well defined on almost all trajectory and does not depend on the starting color.

The result is proved similarly for the drift. $\square$

4 Colored random walks on plain groups

Closed formulas for the drift and entropy of nearest-neighbor random walks on free groups were obtained in [15, 5, 18]. Similar approaches can be carried out for free products of finite groups, of monoids, of finite alphabets [9, 7, 21]. The case of free groups and free products of finite groups is very similar in nature and they can be handled together as done in [14]. As a preliminary step for closed formulas, in this section, we define and study the harmonic measure for colored random walks on plain groups.

Let $G_1, \dots, G_m$ be finite groups and consider the plain group $G = \mathbb{F}_d * G_1 * \cdots * G_m$, with the set of generators

$$S := \bigcup_{i=1}^{d} \{a_i, a_i^{-1}\} \bigcup \left(\bigsqcup_{j=1}^{m} S_j \right),$$

where for all $j = 1, \dots, m$, $S_j := G_j \setminus \{e\}$. Recall the definition of the map Next in (9). Every element $g \in G$ writes uniquely as a word $g = g_1 \cdots g_n$ with $n = |g|$ and $g_i \in S, g_{i+1} \in \mathrm{Next}(g_i)$ for all i. Such words will be called reduced.

If for all $j = 1, \dots, m$, $G_j = \mathbb{Z}/2\mathbb{Z}$, then we say, with a slight abuse of vocabulary, that $G = \mathbb{F}_d * \mathbb{Z}/2\mathbb{Z} * \cdots * \mathbb{Z}/2\mathbb{Z}$ is a free group (two generators are free if there are not inverse).

4.1 The harmonic measure on the boundary

Consider a colored random walk $(X_n)_{n \geq 0}$ on $G \times [r]$ defined by a family (p_g) of matrices. We assume that the walk is quasi-irreducible and nearest-neighbor. We assume furthermore that G is not isomorphic to $\mathbb{Z}$ or $\mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}$. This implies the colored walk is transient, as will be proved in particular in Proposition 9 below.

Define the boundary ∂G as

$$\partial G := \{\xi_0 \xi_1 \cdots \xi_n \cdots, \forall i \geq 0, \xi_i \in G, \xi_{i+1} \in \mathrm{Next}(\xi_i)\}.$$

The multiplication action by G on itself can be extended to ∂G: given $g \in G$, $\xi = \xi_0 \xi_1 \cdots \in \partial G$, define

$$g\xi := \begin{cases} g\xi_0\xi_1\cdots & \text{if } \xi_0 \in \mathrm{Next}(g) \\ (g\xi_0)\xi_1\cdots & \text{if } g\xi_0 \in S \\ \xi_1\xi_2\cdots & \text{if } g = \xi_1^{-1} \end{cases}. \tag{19}$$

The boundary ∂G, which can be seen as a subset of $G^{\mathbb{N}}$, is equipped with the product topology and the corresponding σ-algebra. Given a measure ν on ∂G, let $g \cdot \nu$ be the image measure of ν under the mutiplication by g, that is the measure defined by the fact that

$$\int f(\xi)\, d(g \cdot \nu)(\xi) := \int f(g\xi)\, d\nu(\xi),$$

for all bounded measurable function f on ∂G.

Definition 5. *A colored measure is a family $\nu = (\nu_u)_{u \in [r]}$ of probability measures on ∂G indexed by colors. A colored measure ν is stationary if for all $u \in [r]$,*

$$\nu_u = \sum_{g \in G, v \in [r]} p_g(u, v)\, g \cdot \nu_v. \tag{20}$$

The following result extends Theorem 1.12 in Ledrappier [18] to colored random walks. The proof is exactly the same and is reproduced below.

Proposition 9. *There exists a random variable $X_\infty \in \partial G$ such that X_n converges a.s. to X_∞. The law of X_∞ is called the harmonic measure and is the unique stationary colored measure on the boundary ∂G. It will be denoted $(p_u^\infty)_{u\in[r]}$ where the index u is to be interpreted as the starting color of $(X_n)_{n\geq 0}$.*

Proof. Consider the topology on ∂G defined by the distance

$$d(\xi,\xi') = e^{-|\xi\wedge\xi'|},$$

with $|\xi\wedge\xi'|$ being the length of the prefix common to ξ and ξ'.

This topology makes ∂G a compact set so by the Lévy-Prokhorov theorem, the set $\mathcal{P}(\partial G)$ of probability measures on ∂G is a non-empty convex and compact Hausdorff set when embedded with the weak-* topology. Let $\zeta : \nu_u \mapsto \sum_{g\in G, v\in[r]} p_g(u,v)(g\cdot\nu_v)$. The multiplication in ∂G by an element of G being continuous, ζ is also a continuous map from $\mathcal{P}(\partial G)$ to itself. By the Schauder-Tychonoff theorem, ζ admits a fixed point, which is exactly a stationary measure.

Consider now a stationary measure ν. Since G is not isomorphic to $\mathbb{Z}$ or $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$ the orbits under multiplication are dense in ∂G, that is for all $\xi,\xi'\in\partial G$ and $\epsilon\in(0,1)$ there exists $g\in G$ such that $d(g\xi,\xi')<\epsilon$. From this we deduce that the only closed subsets of ∂G that are invariant under multiplication by G are the empty set and ∂G itself. In particular the support of ν must be ∂G and ν must be continuous: otherwise the set of points with the largest measure would be invariant and finite, so distinct from ∂G.

Stationarity implies that the sequence $(X_n\cdot\nu)$ has the martingale property, in the sense that for every measure set A the sequence $(X_n\cdot\nu(A))$ is a martingale. This martingale is positive and bounded, so $(X_n\cdot\nu)$ converges a.s. and in L^1 to a random measure Z. This implies in particular that a.s. $|X_n|\to\infty$. Indeed, one could otherwise extract a subsequence on which $X_n = g$ for some $g\in G$, with positive probability. Hence $Z = g\cdot\nu$ and because ν is continuous and has support ∂G, this would imply $X_n\to g$ and consequently X_n would be stationary with positive probability. By the quasi-irreducibility assumption, the support of the colored walk is not reduced to e, so Borel-Cantelli lemma shows this is not possible. Note that $|X_n|\to\infty$ shows in particular that X_n is transient.

We prove now that Z must be a Dirac mass. Since $|X_n|\to\infty$ and ν is continuous the measure under $X_n\cdot\nu$ of any cylinder has to tend to 0. Therefore the limit measure $Z(\omega)$ is exactly a Dirac mass on a point $X_\infty(\omega)$. This proves a.s. $X_n\to X_\infty$.

Markov property then shows that the law of X^∞ is indeed stationary. Finally by the martingale property and the L^1 convergence :

$$\nu = \mathbb{E}_e[X_n\nu] \to \mathbb{E}_e[\mathbb{1}_{X_\infty}] = \mathbb{P}_e[X_\infty\in\cdot].$$

Hence ν is also equal to the law X_∞, which proves uniqueness of the stationary measure. $\square$

4.2 Markovian structure of the harmonic measure

A measure on ∂G is uniquely determined by its mass on cylinders. Given a measure ν, we will write $\nu(\xi_1\cdots\xi_n)$ for the mass of the cylinder containing all infinite words which start with the prefix $\xi_1\cdots\xi_n$.

In the colorless case, the tree structure of the group G implies the harmonic measure is Markovian. It can be computed entirely from the solutions of a set of equations derived from the stationarity of the harmonic measure with the following interpretation.

For all $k \geq 0$, let $X_\infty^{(k)}$ be the restriction of X_∞ to the first k-th letters. Thus for all $k \geq 1$ the mass under the harmonic measure of cylinders of size k is given by the law of $X_\infty^{(k)}$. On the other hand $(X_\infty^{(k)})_{k\geq 1}$ is a non-backtracking walk which one can interpret as the loop-erased random walk formed from $(X_n)_{n\geq 0}$. In the standard colorless setting, the tree structure of the group makes the loop-erased random walk a Markov chain whose transition probabilities can be computed easily.

In the colored setting, one can expect to have similar properties but the loop-erased random walk of the process is no longer being a Markov chain and this is not the right process to consider. Instead let $\tau_g := \inf\{n \geq 0, X_n = (g,\cdot)\}$ be the hitting time of g by the random walk X_n and u_k the color at time $\tau_{X_\infty^{(k)}}$. Thus u_k is the first color seen by the random walk X_n when it visits the same element as the k-th step of X_∞, but it is not necessary the color visited at the step X_n where the k-th letter of walk starts coinciding with $X_\infty^{(k)}$ forever. Given $g \in S$, set

$$\mu_g(u,v) := \mathbb{P}_{(e,u)}[X_\infty^{(1)} = g, X_{\tau_g} = (g,v)]. \tag{21}$$

In words, $\mu_g(u,v)$ is the probability that the random walk, starting from (e,u), visits g, with color v for the first time, and later escapes at infinity in direction g. The process considered which is the equivalent of the loop erased random walk in the colored setting is $(X_\infty^{(k)}, u_k)_{k\geq 1}$, which is a colored Markov chain with increment distribution μ.

As we shall check in Lemma 10, the family of matrices $\mu = (\mu_g)_{g\in S}$ is solution of the following set of matrix equations:

$$x_g = p_g \Delta(x)_g + \sum_{\substack{h,h'\in S \\ hh'=g}} p_h x_{h'} + \sum_{h\in \mathrm{Next}(g)} p_{h^{-1}} x_h \Delta(x)_h^{-1} x_g, \quad \forall g \in S, \tag{22}$$

where $\Delta(x)_g$ is the diagonal matrix with entries, for $u \in [r]$,

$$\Delta(x)_g(u,u) := \sum_{h\in \mathrm{Next}(g)} \sum_{v\in[r]} x_h(u,v). \tag{23}$$

In the sequel, $\Delta(\mu)_g$ will be simply written Δ_g.

These equations generalize the so-called traffic equations of Mairesse [14]. Beware that the products are to be understood as matrix products. In particular one has to be careful not to change the order of the different terms when considering products.

Lemma 10. *The family $\mu = (\mu_g)_{g\in S}$ is the unique family of matrices with non-negative entries which sum to a stochastic matrix and which is solution of Equation (22). Moreover, for all starting color $u \in [r]$ and all cylinder $\xi_1 \cdots \xi_n$,*

$$\mathbb{P}_u \left[X_\infty^{(n)} = \xi_1 \cdots \xi_n, u_n = v \right] = \mathbb{1}_u^\top \mu_{\xi_1} \Delta_{\xi_1}^{-1} \mu_{\xi_2} \cdots \Delta_{\xi_{n-1}}^{-1} \mu_{\xi_n} \mathbb{1}_v, \tag{24}$$

and

$$p_u^\infty(\xi_1 \cdots \xi_n) = \mathbb{1}_u^\top \mu_{\xi_1} \Delta_{\xi_1}^{-1} \mu_{\xi_2} \cdots \Delta_{\xi_{n-1}}^{-1} \mu_{\xi_n} \mathbb{1}. \tag{25}$$

Proof. For $n = 1$ equation (24) is a consequence of the definition of the μ_g. Now if $n \geq 2$, the probability that $X_\infty^{(n)} = \xi_1 \cdots \xi_n, u_n = v$ conditional on $X_\infty^{(n-1)} = \xi_1 \cdots \xi_{n-1}, u_{n-1} = u$ is just the probability the random walk escapes at infinity in direction ξ_n, conditionned on the fact it cannot backtrack (in the group), that is

$$\frac{\mu_{\xi_n}(u,v)}{\sum_{h \in \mathrm{Next}(\xi_{n-1})} \sum_{w \in [r]} \mu_h(u,w)} = \Delta_{\xi_{n-1}}^{-1} \mu_{\xi_n}(u,v).$$

An immediate induction yields (24). Summing over colors then gives (25).

Let us now prove now that the matrices μ_g are characterized by Equation (22). Consider any family $(\nu_g)_{g\in S}$ of non-negative matrices, solutions of (22) and such that $\sum_{g\in S} \nu_g$ is a stochastic matrix. Then for all $u \in [r]$ define the measure ν_u^∞ on ∂G by setting for each cylinder $\xi_1 \cdots \xi_n$

$$\nu_u^\infty = \mathbb{1}_u \nu_{\xi_1} \Delta(\nu)_{\xi_1}^{-1} \nu_{\xi_2} \cdots \Delta(\nu)_{\xi_{n-1}}^{-1} \nu_{\xi_n} \mathbb{1}$$

The fact that $\sum_{g\in S} \nu_g$ is a stochastic matrix proves that ν_u^∞ is indeed a probability measure on ∂G.

Let us prove that the hypotheses on ν_g necessarily imply $\nu_u^\infty = p_u^\infty$ for all $u \in [r]$. By uniqueness of the harmonic measure, it suffices to show that ν^∞ is stationary. From (19)-(20), to prove stationarity, we need to show that for all cylinder $\xi_1 \cdots \xi_n$, for all $u \in [r]$,

$$\begin{aligned} \nu_u^\infty(\xi_1 \cdots \xi_n) \quad &= \quad \sum_{v \in [r]} p_{\xi_1}(u,v) \nu_v^\infty(\xi_2 \cdots \xi_n) + \sum_{\substack{g,h \in S \\ gh = \xi_1}} p_g(u,v) \nu_v^\infty(h \xi_2 \cdots \xi_n) \\ &\qquad + \sum_{\substack{g \in S \\ g^{-1}\xi_1 \notin S \cup \{e\}}} p_g(u,v) \nu_v^\infty(g^{-1} \xi_1 \cdots \xi_n). \end{aligned}$$

We use the expression of ν^∞ in this equation and notice there are matrices appearing in both sides of the equation. Furthermore, the case where there exist $g, h \in S$ such that

$gh = \xi_1$ can only occur if g, h and ξ_1 belong to the same finite group G_i, in which case $\mathrm{Next}(h) = \mathrm{Next}(\xi_1)$. Combining these two observations, we find that a sufficient condition for stationarity is

$$\nu_{\xi_1} = p_{\xi_1} \Delta(\nu)_{\xi_1} + \sum_{\substack{g,h \in S \\ gh = \xi_1}} p_g \nu_h + \sum_{\substack{g \in S \\ g^{-1}\xi_1 \notin S \cup \{e\}}} p_g \nu_{g^{-1}} \Delta(\nu)^{-1}_{g^{-1}} \nu_{\xi_1}.$$

This is precisely ensured by the fact that the ν_g are solutions of (22). Indeed, by construction, for $a \in S$, $\mathrm{Next}(a) = \{g \in S : ag \notin S \cup \{e\}\}$ and thus $\{g \in S : g^{-1}\xi_1 \notin S \cup \{e\}\} = \mathrm{Next}(\xi_1^{-1})$.

Thus we just proved that for all $g \in S, u \in [r]$, $p_u^\infty(g) = \sum_{v \in [r]} \mu_g(u,v) = \sum_{v \in [r]} \nu_g(u,v)$. In particular, this yields that $\Delta_g = \Delta(\nu)_g$ for all $g \in S$. Consequently the matrices μ_g and ν_g can be seen as the fixed points of a map $f : M = (M_g)_{g \in S} \mapsto (f(M)_g)_{g \in S}$ from the set of non-negative matrices to itself, where

$$f(M)_g := A_g + \sum_{\substack{h,h' \in S \\ hh' = g}} p_h M_{h'} + \sum_{h \in \mathrm{Next}(g)} p_{h^{-1}} M_h B_g M_g,$$

for all $g \in S$, and A_g, B_g are non-negative matrices that do not depend on M.

One can now argue as in the proof of [14, Lem. 4.7] to deduce that $\mu_g = \nu_g$. Let $\leq$ be the coordinate-wise ordering for matrices: $M \leq N$ if and only if $M(u,v) \leq N(u,v)$ for all $u, v \in [r]$. Consider the matrices m_g defined by

$$m_g(u,v) := \mu_g(u,v) \wedge \nu_g(u,v)$$

As each map $f_g : M \mapsto f(M)_g$ is obviously non-decreasing with respect to the coordinate-wise ordering, one must have $f(m)_g \leq f(\nu)_g = \nu_g$ and $f(m)_g \leq f(\mu)_g = \mu_g$ for all $g \in S$, so $f(m) \leq m$. By Brouwer's fixed point theorem, f must therefore admit another fixed point $m' \leq m$. with the same properties as μ and ν. This m' is thus another solution of (22) and so must satisfy $\sum_{v \in [r]} m'_g(u,v) = \sum_{v \in [r]} \mu_g(u,v)$ for all $u \in [r]$, but since $m' \leq m \leq \mu$ we deduce $\mu = m' = \nu$. □

Hitting probabilities. Let us discuss here another way to compute the harmonic measure through hitting probabilities. It is less direct but we believe it makes the computation more natural in several examples such as free groups.

Given $g \in G, u \in [r]$, recall that τ_g is the hitting time of g. Let $\tau_{(g,u)} := \inf\{n : X_n = (g,u)\}$ be the hitting time of a pair (g,u), so that $\tau_g := \min_u \tau_{(g,u)}$. For $g \in S$, set

$$q_g(u,v) := \mathbb{P}[\tau_g < \infty \text{ and } \tau_g = \tau_{(g,v)}]. \tag{26}$$

Conditioning on X_1 and applying Markov property, one obtains that the family of matrices $q = (q_g)_{g\in S}$ satisfy the matrix equation

$$x_g = p_g + \sum_{\substack{h,h'\in S\\ hh'=g}} p_h x_{h'} + \sum_{h\in \mathrm{Next}(g)} p_{h^{-1}} x_h x_g. \tag{27}$$

In the case where $G = \mathbb{F}_d * \mathbb{Z}/2\mathbb{Z} * \cdots * \mathbb{Z}/2\mathbb{Z}$ is a free group . This equation can again be found for classical nearest-neighbor random walks in [18, 14, 21] but here the equation is a matrix equation.

Lemma 11. *The family $q = (q_g)_{g\in S}$ is the unique solution to Equation (27) among family of matrices $(m_g)_{g\in S}$ such that $\mu_g(u,v) \le m_g(u,v) \le 1$ for all $u,v \in [r]$.*

Proof. One can argue as in the proof of lemma 10 or the proof of [14, Lemma 4.7]: solutions of Equation (27) can be seen as fixed points of quadratic maps that are non-increasing with respect to the coordinate-wise ordering, which cannot have two fixed points. □

The relation between the matrices q_g and μ_g is the following. Suppose the random walk is started at u. In order to have $X_\infty^{(1)} = g, u_1 = v$, the random walk has to visit g with first color v, which occurs with probability $q_g(u,v)$. Arrived at (g,v), it has to escape at infinity in some direction $h \in \mathrm{Next}(g)$. Hence by Markov property,

$$\mu_g(u,v) = q_g(u,v)\Delta_g(v,v), \quad \forall g \in S. \tag{28}$$

This gives yet another way to write equations (24) and (25):

$$\mathbb{P}_u\left[X_\infty^{(k)} = \xi_1\cdots\xi_k, u_k = v\right] = \left(q_{\xi_1}\cdots q_{\xi k-1}\mu_{\xi_k}\right)(u,v), \tag{29}$$

$$p_u^\infty(\xi_1\cdots\xi_k) = \sum_v \left(q_{\xi_1}\cdots q_{\xi k-1}\mu_{\xi_k}\right)(u,v). \tag{30}$$

4.3 Solving the traffic equations

General case. We have seen in the previous section that the harmonic measure is fully described by the family of matrices $\mu = (\mu_g)_{g\in S}$ on $[r]$ defined by (21). By Lemma 10, these matrices are uniquely characterized by the traffic equation (22). It is possible to evaluate numerically μ by iterating the map defining the traffic equations.

The hitting probabilities $q = (q_g)_{g\in S}$ satisfy a simpler quadratic equation (27) than the traffic equation of μ. As explained in the proof of Lemma 11, it is possible to evaluate numerically from above and from below q by iterating the map f defined in (27) such that $q = f(q)$. Moreover, once the hitting probabilities are computed, it is easy to compute μ. Indeed, using (28), the traffic equation reads (21)

$$\mu_g = p_g\Delta_g + \sum_{\substack{h,h'\in S\\ hh'=g}} p_h\mu_{h'} + \sum_{h\in \mathrm{Next}(g)} p_{h^{-1}} q_h \mu_g, \quad \forall g \in S. \tag{31}$$

Observe that $y_g := \Delta_g \mathbb{1} = \mu_g \mathbb{1}$ and $y_g \in \mathbb{R}^r$ is equal to the diagonal of the diagonal matrix Δ_g. By construction, we have that $\sum_g y_g = \mathbb{1}$. If we evaluate the matrix equation (31) on the vector $\mathbb{1}$, we obtain a linear equation for $y = (y_g)_{g\in S}$, seen as a vector with coordinates in $\mathbb{R}^r$, of the form $Ty = 0$ where T is matrix on $S \times S$ with matrix-valued coefficients in $M_r(\mathbb{R})$. Finally, once, Δ_g is known, Equation (31) becomes linear in μ seen as a vector of matrices.

Case of a free group. In some special cases, one can find additional relations allowing to write the matrix μ_g as an explicit function of the matrices q_g. This case occurs for instance when $G = \mathbb{F}_d * \mathbb{Z}/2\mathbb{Z} * \cdots * \mathbb{Z}/2\mathbb{Z}$ is a free group. Given $g \in S$, let d_g be the diagonal matrix with diagonal entries

$$d_g(u,u) := \sum_w q_g(u,w).$$

In words, $d_g(u,u)$ is the probability to ever reach g, starting from the pair (e,u).

Proposition 12. *When $G = \mathbb{F}_d * \mathbb{Z}/2\mathbb{Z} * \cdots * \mathbb{Z}/2\mathbb{Z}$ is a free group, for all $g \in S$ the matrix $(I - q_{g^{-1}}q_g)$ is invertible and*

$$\mu_g(u,v) = q_g(u,v) \sum_{w\in[r]} (I - q_{g^{-1}}q_g)^{-1}(I - d_{g^{-1}})(v,w) \quad \forall g \in S, u,v \in [r]. \tag{32}$$

Proof. In the case of a free group, for all $g \in S, u \in [r]$,

$$\mu_g(u,v) = q_g(u,v)\left(1 - \sum_{w\in[r]} q_{g^{-1}}(v,w) + \sum_{w,z\in[r]} q_{g^{-1}}(v,w)\mu_g(w,z)\right).$$

In particular $q_g(u,v) = 0$ implies $\mu_g(u,v) = 0$. Otherwise, rewrites this equation as

$$\frac{\mu_g(u,v)}{q_g(u,v)} = 1 - d_{g^{-1}}(v,v) + \sum_{w,z\in[r]} q_{g^{-1}}(v,w)q_g(w,z)\frac{\mu_g(w,z)}{q_g(w,z)}.$$

The right-hand being independant of u, so is the left-hand side. Consequently let $x_v := \mu_g(u,v)/q_g(u,v)$. Then one can again rewrite the above equation as the matrix linear equation

$$x = (I - d_{g^{-1}})\mathbb{1} + q_{g^{-1}}q_g x$$

Provided $I - q_{g^{-1}}q_g$ is invertible, solving this equation yields the desired expression for $\mu_g(u,v)$.

Therefore we are left to prove that $I - q_{g^{-1}}q_g$ is invertible. This can be justified by the fact that $\sum_{v\in[r]}\sum_{n\geq 0}(q_{g^{-1}}q_g)^n(u,v)$ is the average number of times the walk goes to g^{-1} and comes back to e when starting at color u. By transience of the walk, this number must be finite. Hence the sum $\sum_{n\geq 0}(q_{g^{-1}}q_g)^n$ is convergent and the inverse of this matrix is precisely $I - q_{g^{-1}}q_g$. □

We now further study the matrix equation (26) satisfied by the hitting probabilities $(q_g)_{g\in S}$ in the case where $G = \mathbb{F}_d * \mathbb{Z}/2\mathbb{Z} * \cdots * \mathbb{Z}/2\mathbb{Z}$ is a free group. As already pointed, in this case, we have for all $g \in S$,

$$q_g = p_g + \sum_{h \neq g^{-1}} p_{h^{-1}} q_h q_g. \tag{33}$$

In the colorless case, it is possible to reduce Equation (33) to a scalar equation, see for example [18]. We extend this computation to the colored case. We define the matrix in $M_r(\mathbb{R})$,

$$z = I - \sum_{g\in S} p_{g^{-1}} q_g. \tag{34}$$

We now express q_g as a function of $(z, p_g, p_{g^{-1}})$ and find a closed equation satisfied by z. For simplicity, we assume that for all $g \in G$, the matrix p_g is invertible. If this is not the case, a similar argument holds but one should be careful with pseudo-inverses.

We define the matrices in $M_{2r}(\mathbb{R})$:

$$P_g = \begin{pmatrix} 0 & p_g \\ p_{g^{-1}} & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} z & 0 \\ 0 & z \end{pmatrix} \quad \text{and} \quad Q_g = \begin{pmatrix} 0 & q_g \\ q_{g^{-1}} & 0 \end{pmatrix}.$$

From our assumption, P_g is invertible and

$$P_g^{-1} = \begin{pmatrix} 0 & (p_{g^{-1}})^{-1} \\ p_g^{-1} & 0 \end{pmatrix}.$$

Applying (33) to g and g^{-1}, we find that, for all $g \in S$, $zq_g = p_g - p_g q_{g^{-1}} q_g$ and thus

$$ZQ_g = P_g - P_g Q_g^2.$$

We set $Z_g = P_g^{-1} Z$. The above equation rewrites

$$(Z_g + Q_g) Q_g = I. \tag{35}$$

In particular $Z_g = Q_g^{-1} - Q_g$ so Z_g and Q_g commute. We may thus solve the quadratic equation $Q_g^2 + Z_g Q_g - I = 0$ with Q_g as unknown as in the scalar case. Completing the square yields

$$(Q_g + Z_g/2)^2 = I + Z_g^2/4.$$

Therefore, for some proper choice of the matrix square root function, we get

$$Q_g = \frac{1}{2}\left(\sqrt[*]{4I + (P_g^{-1} Z)^2} - P_g^{-1} Z \right), \tag{36}$$

where $\sqrt[*]{\cdot}$ is a notation to stress that the choice of the square root is unknown. First, as Z_g and Q_g commute, the eigenvalues completely determine the square root. Also, since Q_g is a block antidiagonal matrix, for every of its eigenvalue λ, $-\lambda$ must also be an eigenvalue with the same multiplicity (algebraic and geometric). In particular, we are left with at most 2^r choices for the square root to pick in (36) (a choice of sign for each eigenvalue pair $(\lambda, -\lambda)$).

There is one useful property to further determine the square root. Consider

$$R_g = Q_g P_g^{-1} = \begin{pmatrix} q_g p_g^{-1} & 0 \\ 0 & q_{g^{-1}} p_{g^{-1}}^{-1} \end{pmatrix}.$$

We observe from (33) that $q_g p_g^{-1} = (I - \sum_{h \neq g^{-1}} p_{h^{-1}} q_h)^{-1}$. The matrix $\sum_{h \neq g^{-1}} p_{h^{-1}} q_h$ is substochastic: it has non-negative entries and the sum over each row is less or equal than one (from a starting color, it corresponds to the probability that the colored walk killed when visiting g comes back to e after some time). In particular, the matrix $\sum_{h \neq g^{-1}} p_{h^{-1}} q_h$ has spectral radius less than one. Thus all eigenvalues of $q_g p_g^{-1}$ and R_g have positive real part. Using (34), the same argument shows that z and Z have all their eigenvalues with positive real parts.

Finally, from Equation (34), we have $Z = I - \sum_g P_g Q_g = I - \sum_g P_g R_g P_g$. We deduce that

$$Z = I - \frac{1}{2} \sum_{g \in S} \left(P_g \sqrt[*]{4I + (P_g^{-1} Z)^2} - Z \right). \tag{37}$$

Up to this issue of square root, we thus have found a fixed point equation satisfied by z (in Equation (37)) and expressed q_g as a function of $(z, p_g, p_{g^{-1}})$ (in Equation (36)). If $r = 1$, we can retrieve a known formula for colorless random walks. We note that Equation (37) should be compared to Proposition 3.1 in Lehner [19] where a related formula is derived in a self-adjoint case (it can be checked that $2z$ is the inverse of the diagonal term of the Green function $(I - \tilde{\mathcal{P}})^{-1}(e,e)$ where $(I - \tilde{\mathcal{P}})^{-1}$ is seen as an infinite matrix indexed by $G \times G$ with coefficients in $M_r(\mathbb{C})$).

5 Computing drift and entropy

5.1 Computation of the drift: proof of Theorem 4

A proof of Theorem 4 in the colorless case is given in [18] for the free group. The proof can be adapted for more general free products, as mentionned in [14, 21]. It is also presented in a slightly different way in [9]. This is the latter proof that we here extend in the colored case.

Proof of Theorem 4. We aim to compute the limit of $\mathbb{E}|X_n|/n$. This can be seen the Cesàro limit of the sequence $(\mathbb{E}|X_{n+1}| - \mathbb{E}|X_n|)_{n \geq 0}$. Here we deliberately omitted to precise the starting color, as the final result does not depend on it by Proposition 2. In particular one can choose the starting color to be distributed according to the measure π. By Markov property and the invariance of π, for all $n \geq 0$

$$\begin{aligned}
\mathbb{E}_\pi |X_{n+1}| - \mathbb{E}_\pi |X_n| &= \sum_{u,v\in[r]} \pi(u)p_g(u,v)\mathbb{E}_v |gX_n| - \sum_u \pi(u)\mathbb{E}_u |X_n| \\
&= \sum_{u,v\in[r]} \pi(u)p_g(u,v) \left(\mathbb{E}_v |gX_n| - \mathbb{E}_v |X_n|\right) \\
&= \sum_{u,v\in[r]} \pi(u)p_g(u,v) \left(\mathbb{E}_v \left[|gX_n| - |X_n|\right]\right).
\end{aligned}$$

We fix a color v. The random variable $Z_n \coloneqq |gX_n| - |X_n|$ takes values in $\{-1,0,1\}$. On the other hand, X_n converges to the random variable X_∞ as n goes to infinity. Consequently gX_n converges to $gX_\infty \in \partial G$ and Z_n converges to a random variable $Z_\infty \in \{-1,0,1\}$. It is now easy to see what happens:

$$Z_\infty = \begin{cases} -1 & \text{if } X_\infty^{(1)} = g^{-1} \\ 1 & \text{if } X_\infty^{(1)} \in \mathrm{Next}(g) \\ 0 & \text{otherwise} \end{cases}.$$

Hence, by the dominated convergence theorem, we find

$$\begin{aligned}
\mathbb{E}_v [Z_n] \to \mathbb{E}_v [Z_\infty] &= \int Z_\infty \, dp_v^\infty \\
&= -\sum_{w\in[r]} \mu_{g^{-1}}(v,w) + \sum_{h\in\mathrm{Next}(g)} \sum_{w\in[r]} \mu_h(v,w),
\end{aligned}$$

as requested. □

5.2 Entropy: : proofs of Theorem 5 and Theorem 6

The following computation is originally due to Kaimanovich and Vershik [15] and is here adapted to the matrix case. Given a discrete random variable X and a σ-algebra $\mathcal{F}$, the conditional entropy of X with respect to $\mathcal{F}$ is defined as

$$H(X \mid \mathcal{F}) = \mathbb{E}\left[-\sum_x \mathbb{P}(X = x \mid \mathcal{F}) \log \mathbb{P}(X = x \mid \mathcal{F})\right].$$

Let $\mathcal{T} \coloneqq \bigcap_{n\geq 0} \sigma\{X_k, k \geq n\}$ be the tail σ-algebra. As one can check easily, the tail σ-algebra is also the σ-algebra generated by X_∞. Therefore a function is $\mathcal{T}$-measurable if and only if it writes as a function of X_∞.

Proposition 13. *The entropy of a colored random walk satisfies*

$$h = H_\pi(X_1) - H_\pi(X_1 \mid \mathcal{T}). \tag{38}$$

Proof. By the Markov property for all $n \geq 1$

$$H_\pi(X_1 \mid X_n, X_{n+1}, \dots) = H_\pi(X_1 \mid X_n)$$

When $n \to \infty$ the left-hand side converges to $H_\pi(X_1|\mathcal{T})$. On the other hand,

$$\begin{aligned} H_\pi(X_1 \mid X_n) &= H_\pi(X_1, X_n) - H_\pi(X_n) \\ &= H_\pi(X_1) + H_\pi(X_n \mid X_1) - H_\pi(X_n) \\ &= H_\pi(X_1) - (H_\pi(X_n) - H_\pi(X_{n-1})) \,. \end{aligned}$$

In the last line we have used equation (18). Since $H_\pi(X_n)/n$ converges to h as n goes to infinity, one must also have that $H_\pi(X_n) - H_\pi(X_{n-1})$ converges to h, hence the result. □

Given a σ-algebra $\mathcal{F}$ and $g, h \in G, u, v \in [r]$, let $\mathbb{P}[(g,u),(h,v) \mid \mathcal{F}]$ be the conditional transition probability with respect to $\mathcal{F}$, that is

$$\mathbb{P}[(g,u),(h,v) \mid \mathcal{F}] \coloneqq \mathbb{P}_{g,v}[X_1 = (h,v) \mid \mathcal{F}] \,.$$

The stationarity of the harmonic measure

$$p_u^\infty = \sum_{\substack{g \in G \\ v \in [r]}} p_g(u,v)\, g \cdot p_v^\infty \tag{39}$$

implies in particular that the measures $g \cdot p_v^\infty$ are absolutely continous with respect to p_u^∞ if $p_g(u,v) > 0$.

Proposition 14. *The conditional probabilities with respect to $\mathcal{T}$ are given by:*

$$\mathbb{P}[(g,u),(h,v) \mid \mathcal{T}] = p_{g^{-1}h}(u,v) \frac{d\, hp_v^\infty}{d\, gp_u^\infty} \quad a.s.. \tag{40}$$

Sketch of proof. The result is true when replacing $\mathcal{T}$ with the invariant σ-algebra $\mathcal{I}$ which is the σ-algebra generated by all bounded invariant functions, that is, all bounded functions $f : (G \times [r])^{\mathbb{N}} \to \mathbb{R}$ such that $f(x_0, x_1, \dots) = f(x_{1,}, x_2, \dots)$. This is the content of [20, Lemma 14.33].

Generally, the invariant and tail σ-algebras do not coincide but they do for transitive Markov chains, up to negligible sets. The proof goes basically like this: there is an equivalence between the set of bounded invariant functions and the set of bounded harmonic functions. There is a similar equivalence between bounded tail functions and bounded space-time harmonic functions. Using a zero-two law, one then shows that space-time harmonic functions do not actually depend on the time parameter and hence are classical harmonic functions. The equivalences between functions thus prove the equivalence of the σ-algebras. See [20, Section 14.6] for details.

For colored Markov chains, the arguments extend naturally, using quasi-irreducibility and quasi-transitivity. Therefore the above formula is correct with $\mathcal{T}$ instead of $\mathcal{I}$. □

Proof of Theorem 5. By Proposition 13 and Equation (39),

$$\begin{aligned}
h &= H_\pi(X_1) - H_\pi(X_1 \mid \mathcal{T}) = H_\pi(X_1) + \sum_u \pi(u) \int \log \mathbb{P}_{(e,u)}[X_1 \mid \mathcal{T}](X_\infty)\, d\mathbb{P}_u \\
&= H_\pi(X_1) + \sum_u \pi(u) \sum_{\substack{g\in S \\ v\in[r]}} \int p_g(u,v) \log \mathbb{P}_{(e,u)}[X_1 = (g,v) \mid \mathcal{T}](g\xi)\, dp_v^\infty(\xi) \\
&= H_\pi(X_1) + \sum_{u\in[r]} \pi(u) \sum_{\substack{g\in S \\ v\in[r]}} p_g(u,v) \int \log \mathbb{P}_{(g^{-1},u)}[X_1 = (e,v) \mid \mathcal{T}](\xi)\, dp_v^\infty(\xi)
\end{aligned}$$

Now, using Proposition 14, we find

$$\begin{aligned}
h &= H_\pi(X_1) + \sum_u \pi(u) \sum_{g\in\mathbb{F}_d, v\in[r]} p_g(u,v) \int \log\left(p_g(u,v) \frac{dp_v^\infty}{dg^{-1}p_u^\infty}(\xi)\right) dp_v^\infty(\xi) \\
&= -\sum_u \pi(u) \sum_{g\in\mathbb{F}_d, v\in[r]} p_g(u,v) \int \log\left(\frac{dg^{-1}p_u^\infty}{dp_v^\infty}(\xi)\right) dp_v^\infty(\xi).
\end{aligned}$$

It concludes the proof. □

Equation (10) is the matrix version of the known formula for colorless random walks [15, 18]. It shows that the computation of the entropy ultimately comes down to the computation of some Radon-Nikodym derivatives. In the colorless case, this computation goes basically as follows.

Consider a cylinder $\xi_1 \cdots \xi_n$. Compute $g \cdot p^\infty(\xi_1 \cdots \xi_n)$ distinguishing the different cases occurring:

- if $g = \xi_1$ then $g \cdot p^\infty(\xi_1 \cdots \xi_n) = p^\infty(\xi_2 \cdots \xi_n)$
- if $\xi_1 \in G_i, g \neq \xi_1$ and there exists $h \in G_i$ such that $gh = \xi_1$, then $g \cdot p^\infty(\xi_1 \cdots \xi_n) = p^\infty(h\xi_2 \cdots \xi_n)$
- otherwise, $g \cdot p^\infty(\xi_1 \cdots \xi_n) = p^\infty(g^{-1}\xi_1 \cdots \xi_n)$.

As one can check, the last case occurs if and only if $\xi_1 \in \mathrm{Next}(g^{-1})$. Expand now the expressions above into products using from (30) that $p^\infty(\xi_1 \cdots \xi_n) = q_{\xi_1} \cdots q_{\xi_{n-1}} \mu_{\xi_n}$ to get

$$\frac{g \cdot p^\infty(\xi_1 \cdots \xi_n)}{p^\infty(\xi_1 \cdots \xi_n)} = \begin{cases} 1/q_g & \text{if } \xi_1 = g, \\ q_{g^{-1}} & \text{if } \xi_1 \in \mathrm{Next}(g^{-1}), \\ q_h/q_{\xi_1} = q_{g^{-1}\xi_1}/q_{\xi_1} & \text{if } h \in S \text{ and } gh = \xi_1. \end{cases}$$

As will be proved after, the left-hand side converges to the Radon-Nikodym derivative as $n \to \infty$. One can then easily express the integral (10) in terms of the scalars p_g, q_g and μ_g and obtain the formulas given in [18, 21].

For colored random walks, the same computations can be made with matrices except that the cancellation between products no longer takes place. Instead, one has to deal with an infinite product of random matrices. For example, as one might guess from (19)-(30), it is true that

$$\frac{dgp_v^\infty}{dp_u^\infty}(\xi) = \lim_{n\to\infty} \frac{\mathbb{1}_v^\top\, q_{g^{-1}} q_{\xi 1} \cdots q_{\xi n-1} \mu_{\xi n} \mathbb{1}}{\mathbb{1}_u^\top\, q_{\xi 1} \cdots q_{\xi n-1} \mu_{\xi n} \mathbb{1}} \quad \text{a.s.,} \tag{41}$$

if $\xi_1 \in \mathrm{Next}(g^{-1})$. In the subsection, we prove that this kind of convergence holds and that the limit can be expressed thanks to random probability measure on $[r]$ which is uniquely characterized by an invariance property.

5.3 Convergence in direction for inhomogeneous products of matrices

In this final subsection, we prove Theorem 6. The computation of the Radon-Nikodym derivatives in the integral formula of the entropy requires to investigate infinite products of random, non-negative matrices (that is, with non-negative entries). We use first use results for non-negative, deterministic matrices [24] to justify limits of infinite matrix products. This yields formulas like (41) and a first expression for the entropy (45). Then we use results for products of random matrices [3, 4] to describe the law of the limits obtained, which yields eventually Proposition 6.

Convergence of the Radon-Nikodym derivatives. Instead of the general results in [24], it turns out that we can directly use a result from Lalley who already had to inquire about infinite matrix products in his study of finite range random walks [16]. It does not seem however that his methods can be interpreted as an application of the linearization trick.

Let X be a finite set and Y a subset of $X \times X$ such that for all $x \in X$, the set $\{y \in X, (x,y) \in T\}$ is non empty. Let $\Sigma := X^{\mathbb{Z}}$ be the space of doubly infinite sequences $(\xi_n)_{n\in\mathbb{Z}}$ with values in X, such that $(\xi_n, \xi_{n+1}) \in Y$ for all $n \in \mathbb{Z}$. Let $\sigma : (\xi_n)_{n\in\mathbb{Z}} \mapsto (\xi_{n+1})_{n\in\mathbb{Z}}$ denote the standard shift on Σ and given a function $f : \Sigma \to \mathbb{C}$ and $n \in \mathbb{N}$, write

$$S_n f := f + f\circ\sigma + \dots f\circ\sigma^{n-1}.$$

Recall the definition of $\mathcal{P}^+$ in (12).

Proposition 15 ([16, Prop 5.2])**.** *Let $(M_x)_{x\in X}$ be a family of $r\times r$ matrices with non-negative entries. Assume there exists integers $m \geq 0, k \geq 1$ and a function $B : X^k \to 2^{[r]} \setminus \{\varnothing\}$ (the set of non-empty subsets of $[r]$), such that for every $n \geq m$ and every family $x_1, \dots, x_{n+k} \in X$ with $(x_i, x_{i+1}) \in Y$ for all i,*

$$(M_{x_1} M_{x_2} \cdots M_{x_{n+k}})_{u,v} > 0 \Leftrightarrow v \in B(x_{n+1}, \dots, x_{n+k}) \qquad \forall u, v \in [r] \tag{42}$$

Then there exist constants $C > 0$ and $0 < \alpha < 1$, maps $\varphi, \gamma : \Sigma \to \mathbb{R}$ and $V, W : \Sigma \to \mathcal{P}^+$ such that for all $\xi \in \Sigma$,

$$\left\| e^{-S_n\varphi(\xi)} M_{\xi_1} M_{\xi_2} \cdots M_{\xi_n} - \gamma(\sigma_n \xi) V(\xi) W(\sigma^n \xi)^\top \right\| \leq C\alpha^n, \tag{43}$$

where $V = V(\xi_1, \xi_2, \dots)$ depends only on the "forward coordinates", while $W = W(\xi_0, \xi_{-1}, \dots)$ depends only on the "backward coordinates". Furthermore,

$$M_{\xi_1} V(\sigma\xi) = e^{\varphi(\xi)} V(\xi). \tag{44}$$

Proposition 15 states shows under condition (42), the product $M_{\xi_1} M_{\xi_2} \cdots M_{\xi_n}$ tends up to renormalization factor to a rank one matrix whose range is spanned by the vector $V(\xi)$. This result is well known for powers of a matrix with positive entries, in which case the vector V is nothing but the Perron-Frobenius eigenvector. Proposition 15 is thus a generalization of the Perron Frobenius theory to inhomogeneous products of non-negative matrices with possibly zero columns (condition (42)). We refer to [16] for additional properties of the functions φ, γ with the vectors V, W which will not be needed here.

To apply Proposition 15 in context of linearized random walks, take $X = S$ and $(x, y) \in Y$ if and only if $y \in \mathrm{Next}(g)$.

Corollary 16. *Suppose the matrices $(q_g)_{g \in S}$ satisfy the hypothesis of Proposition 15 and let $V : \partial G \to \mathcal{P}^+$ be the corresponding map. Then for all $u, v \in [r]$, $\xi \in \partial G$,*

$$\frac{dgp_v^\infty}{dp_u^\infty}(\xi) = \begin{cases} \frac{\langle \mathbb{1}_v \,,\, V(\sigma\xi) \rangle}{\langle q_g(u,\cdot) \,,\, V(\sigma\xi) \rangle} & \text{if } \xi_1 = g, \\ \frac{\langle q_{g^{-1}}(v,\cdot), V(\xi) \rangle}{\langle \mathbb{1}_u \,,\, V(\xi) \rangle} & \text{if } \xi_1 \in \mathrm{Next}(g^{-1}), \\ \frac{\langle q_h(v,\cdot) \,,\, V(\sigma\xi) \rangle}{\langle q_{gh}(u,\cdot) \,,\, V(\sigma\xi) \rangle} & \text{if } h \in S \text{ and } gh = \xi_1. \end{cases} \tag{45}$$

Proof. From (19)-(30), for any cylinder $\xi_1 \cdots \xi_n$

$$g \cdot p_v^\infty(\xi_1 \cdots \xi_n) = \begin{cases} \mathbb{1}_v^\top q_{\xi_2} \cdots q_{\xi_{n-1}} \mu_{\xi_n} \mathbb{1} & \text{if } \xi_1 = g, \\ \mathbb{1}_v^\top q_{g^{-1}} q_{\xi_1} \cdots q_{\xi_{n-1}} \mu_{\xi_n} \mathbb{1} & \text{if } \xi_1 \in \mathrm{Next}(g^{-1}), \\ \mathbb{1}_v^\top q_h q_{\xi_2} \cdots \mu_{\xi_n} \mathbb{1} & \text{if } h \in S \text{ and } gh = \xi_1. \end{cases}$$

Let $\xi \in \partial G$. Combining the previous computation with Proposition 15, there exist $\lambda_n > 0$, uniformly lower bounded, and $\alpha < 1$ such that if $\xi_1 = g$,

$$\begin{aligned} \frac{g \cdot p_v^\infty(\xi_1 \cdots \xi_n)}{p_u^\infty(\xi_1 \cdots \xi_n)} &= \frac{\mathbb{1}_v^\top \left(q_{\xi_2} \cdots q_{\xi_{n-1}} \right) \left(\mu_{\xi_n} \mathbb{1} \right)}{q_g(u,\cdot)^\top \left(q_{\xi_2} \cdots q_{\xi_{n-1}} \right) \left(\mu_{\xi_n} \mathbb{1} \right)} \\ &= \frac{\mathbb{1}_v^\top V(\sigma\xi)\lambda_n + O(\alpha^n)}{q_g(u,\cdot)^\top V(\sigma\xi)\lambda_n + O(\alpha^n)} \\ &\xrightarrow[n\to\infty]{} \frac{\langle \mathbb{1}_v \,,\, V(\sigma\xi) \rangle}{\langle q_g(u,\cdot) \,,\, V(\sigma\xi) \rangle}. \end{aligned}$$

The other cases are treated similarly.

On the other hand

$$\frac{g\cdot p_v^\infty(\xi_1\cdots\xi_n)}{p_u^\infty(\xi_1\cdots\xi_n)} = \int_{\partial G}\frac{dgp_v^\infty}{p_u^\infty}\frac{\mathbb{1}_{\xi_1\cdots\xi_n}}{p_u^\infty(\xi_1\cdots\xi_n)}\,dp_u^\infty$$

but as $n\to\infty$ the measure $\frac{\mathbb{1}_{\xi_1\cdots\xi_n}}{p_u^\infty(\xi_1\cdots\xi_n)}\,dp_u^\infty$ converges to the Dirac mass at ξ: for all integer k and $\epsilon>0$, the mass of all k-cylinders except the cylinder $\xi_1\cdots\,\xi_k$ is bounded by ϵ for n large enough. Thus the above limits give indeed the Radon-Nikodym derivative. □

The next proposition asserts that for the colored random walks that we are mainly interested in, Corollary 16 applies.

Proposition 17. *Suppose the colored random walk $(X_n)_{n\geq 0}$ is a linearized random walk obtained from a finite range random walk $(Y_n)_{n\geq 0}$ via one of the procedures presented in Section 2. If $(Y_n)_{n\geq 0}$ is irreducible, then the matrices $(q_g)_{g\in S}$, satisfy the hypothesis of Proposition 15.*

Proof. Beware that the notations are the contrary to those of Section 2: here $(X_n)_{n\geq 0}$ is the colored walk linearizing the finite range walk $(Y_n)_{n\geq 0}$.

Let us first review briefly a relevant argument in [16]. Lalley proves that the hypothesis of Proposition 15 is satisfied by matrices H_g which are related to the walk $(Y_n)_{n\geq 0}$. These matrices are defined as follows. Letting K denote the support of the finite range random walk $(Y_n)_{n\geq 0}$, define $L := \max\{|g|, g\in K\}$ and $\mathcal{B} := \{g\in G, |g|\leq L\}$. The matrices H_g will be taken in $\mathbb{R}^{\mathcal{B}\times\mathcal{B}}$. Given $g\in G$, define

$$T_g := \inf\{n\geq 0, Y_n\in g\mathcal{B}\},$$

that is T_g is the first time the representative of Y_n writes as g times an element in $\mathcal{B}$. When G is a plain group, representatives are unique so T_g is well defined. For all $a,b\in\mathcal{B}$, the (a,b)-entry of H_g is then defined as

$$H_g(a,b) := \mathbb{P}\left[T_g<\infty, Y_{T_g}=gb\,|\,Y_0=a\right]$$

In [16], Lalley only considers the case of a free group. The following properties can however be easily extended to plain groups.

The matrices H_g, $g\in S$ have the property that if $g=g_1\cdots\,g_n$, then $H_g=H_{g_1}\cdots H_{g_n}$, see [16, Prop 2.3]. Thus proving that the family $(H_g)_{g\in S}$ satisfies condition (42) is equivalent to prove that there exists $k\geq 1$ and $B: S^k\to 2^{\mathcal{B}}\setminus\{\varnothing\}$ such that for all $g\in G$ with length $|g|$ large enough,

$$H_g(a,b)>0 \Leftrightarrow b\in B \tag{46}$$

where B is a subset of $\mathcal{B}$ that depends only on the last k letters of g. This is exactly what is done in [16, Proposition 5.3], under the condition that $(Y_n)_{n\geq 0}$ is irreducible.

We now come back to the proof of Proposition 17. As for the matrix H_g, if g writes as a reduced word $g = g_1 \cdots g_n$, then we set $q_g = q_{g_1} \cdots q_{g_n}$, so one can argue as for the H_g, namely prove that $q_g(u,v) > 0$ for $|g|$ large enough if and only if v is in some subset of $[r]$ that depends only on the last letters of g.

The two constructions of the linearized random walks considered are made so that if $Y_0 = e$ and $X_0 = (e,1)$, one can couple $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ to have $Y_n = X_{\tau_n}$ where τ_n always corresponds to a return time at color 1.

Moreover, colors are associated to elements of K: for all $v \in [r], v \neq 1$, the colored walk X_n, when at color 1, needs to be multiplied by specific elements of K in order to pass through the color v. More precisely there exists (at least one) $h \in K$ with representative $h = h_1 \cdots h_k h_{k+1} \cdots h_l$, with the property that when at color 1, the random walk (X_n) reaches color v when being multiplied by $h_1 \cdots h_k$. The main difference about the two constructions is the number of such h that one can associate to the color v, but this has no consequence for what follows.

Let $K(v)$ be the set of such elements h. For each $h = h_1 \cdots h_k h_{k+1} \cdots h_l \in K(v)$, set $p(h) := h_1 \cdots h_k$ and $s(h) := h_{k+1} \cdots h_l$. By what precedes, to have color v when visiting g, X_n has first to get to $(gp(h)^{-1}, 1)$ for some $h \in K(v)$, and then be multiplied by $p(h)$. Arrived at g, it has no other choice then than to be multiplied by $s(h)$.

The same goes for the starting color u: if the walk is at (e,u), then it is bound to arrive at $(s(h'), 1)$ for some $h' \in K(u)$.

Thus using that $X_{\tau_n} = Y_n$, one deduces that the colored random walk can go from (e,u) to (g,v) if and only if the finite range random walk can go from one of the $s(h'), h' \in K(u)$ to $gp(h)^{-1}$ for some $h \in K(v)$ and then to $gs(h)$. Hence for any $h \in K(v), h' \in K(u)$,

$$q_g(u,v) \geq H_g(s(h'), s(h)).$$

Provided $|g|$ is large enough, Equation (46) precisely states that the right-hand side is positive if and only if $s(h)$ is in some subset of $\mathcal{B}$ that only depends on the last letters of g. Consequently the positivity of $q_g(u,v)$ does not depend on u for $|g|$ large enough, which proves the result. □

Law of Radon-Nikodym derivatives. The previous result is purely deterministic and does not take into account that the process $(X_\infty^{(n)}, u_n)_{n\geq 0}$ is a Markov chain. Our goal in this paragraph is to determine the law of the probability vector V appearing in Corollary 16. The monograph [3] (see also [4]) state results for products of iid random matrices which, as we shall see, apply to Markovian products as well.

Remploying the general framework of Proposition 15, consider a Markov chain $(Z_n)_{n\geq 0}$ on the finite state space X with transition matrix Q and $Y \coloneqq \{(x,y) \in X \times X, Q(x,y) > 0\}$. Let $(M_x)_{x\in X}$ be a family of $r \times r$ non-negative matrices, such that condition (42) holds.

Consider the sequence defined by $Y_n \coloneqq M_{Z_n}, n \geq 1$. By Proposition 15, a.s. the product $Y_1 \cdots Y_n$ is non-zero and converges in direction to a rank one matrix spanned by a random vector $V \in \mathcal{P}^+$. If one multiplies the product on the left by another matrix M_x, one still has a products of matrices M_x, so one can expect the law of V to satisfy some invariance property.

For all non-negative matrix $A \in M_r(\mathbb{R})$ and all $z \in \mathcal{P}^+$, we define Az to be the normalized image of z (identified with a vector in $\mathbb{R}^r$) that makes it a probability vector, provided $Az \neq 0$. The latter case will not be an issue: condition (42) implies in particular that matrices M_x have no zero row, and vectors of $\mathcal{P}^+$ have no zero coordinate so $M_x z$ is well defined for all $x \in X, z \in \mathcal{P}^+$.

Since Z_n is a Markov chain, the law of V may depend on the starting state of Z_n. Hence it is natural to think of it as a colored measure (colors being here the states $x \in X$).

Definition 6. *A family $\nu = (\nu_x)_{x\in X}$ of probability measure on $\mathcal{P}^+$ is called a colored measure. Let $Q * \nu$ denote the colored measure defined by*

$$\int f(z)\, d(Q * \nu)_x(z) = \sum_y \int f(M_y z) Q(x,y)\, d\nu_y(z)$$

for all bounded measurable function f on $\mathcal{P}^+$ and $x \in X$. The colored measure ν is said to be invariant if $Q * \nu = \nu$.

Lemma 18. *Let ν_x be the law of V when Z_n is started at x, the colored measure $(\nu_x)_{x\in X}$ is the unique colored measure which is invariant with respect to Q.*

Proof. We use Markov property together with (44) to obtain the invariance of the law of V.

The proof of uniqueness is similar to the proof for the harmonic measure in Proposition 9. Consider another invariant colored measure $(\rho_x)_{x\in X}$. By invariance, for all bounded measurable function f on $\mathcal{P}^+$ the sequence

$$M_n \coloneqq \int f(Y_1 \cdots Y_n w)\, d\rho_{Z_n}(w), \quad n \geq 0$$

is a bounded martingale with respect to the filtration $\mathcal{F}_n \coloneqq \sigma(Z_0, Z_1, \ldots, Z_n)$, so M_n converges a.s. and in L^1. This being true for all bounded measurable function, this implies the measures $Y_1 \cdots Y_n * \rho_{Z_n}$ converges weakly to a measure. Now, because the product $Y_1 \cdots Y_n$ converges in direction to V, the limit is necessarily the Dirac mass at V.

On the other hand, the martingale property with the L^1 convergence gives that for all bounded measurable functions f on $\mathcal{P}^+$, for all $x \in X$,

$$\mathbb{E}_x\left[f(V)\right] = \mathbb{E}_x\left[\int f\, \delta_V\right] = \int f\, d\rho_x.$$

Therefore ρ_x must be the law of V when $Z_0 = x$, that is $\rho_x = \nu_x$, and is consequently unique. □

We are finally ready to prove Theorem 6.

Proof of Theorem 6. We apply Lemma 18 in the context of colored random walks, with $X = S \times [r]$ and $Q((g,u)(h,v)) = \mu_h(u,v)$. Due to the form of the transition probabilities the law of V is only indexed by colors. It remains to use Corollary 16 in conjonction with the integral formula of the entropy given by Theorem 5. □

6 Application

We now give an example of how the linearization trick can be used in order to compute the drift of a random walk. The example is simple enough so that the linearization trick is actually not needed. However it is a good illustration of its usefulness: whereas the direct computation requires some thought and can appear tedious, the linearization trick allows to hide all the technical aspects into matrix multiplications. The computation in the end is the same, but it can be handled quite smoothly with the linearization trick.

Let G_1, G_2 be finite groups and consider the group $G = G_1 * G_2$ with the usual set of generators $S = G_1 \bigsqcup G_2 \smallsetminus \{e\}$. Elements g of G can be written as words of length $|g|$ with letters alternatively taking values in G_1 or G_2.

Let k_i denote the cardinality of $G_i \smallsetminus \{e\}$, $i = 1,2$, $L \leq 1$ be a fixed integer and consider the following random walk $(X_t)_{t\geq 0}$ supported on the set of words of length smaller than L. At each step the walk is multiplied on the right by an element chosen uniformly among words of length $n \leq L$ that start with $i \in \{1,2\}$ with probability $p_i(n)$. In other words for all $g = g_{i_1} \cdots g_{i_n}$ of length $n \leq L$, with $i_1, \ldots, i_n \in \{1,2\}, i_1 \neq i_2 \cdots \neq i_n$,

$$p_g = \frac{p_{i_1}(n)}{k_{i_1} \cdots k_{i_n}},$$

where $p = (p_i(n))_{n\leq L, i=1,2}$ is a probability vector : $\sum_{i,n} p_i(n) = 1$. Of course in the denumerator above, one could regroup powers of k_1, k_2 together. This random walk is the generalization of an example given in [21] for which the drift is explicitly computed in the nearest-neighbor case $L = 1$.

As seen in the proof of Theorem 4, the drift is given by

$$\gamma = \lim_{t\to\infty} \frac{1}{t} \sum_{k=0}^{t-1} \mathbb{E}\left[|X_{k+1}| - |X_k|\right].$$

In this model, the knowledge of what group the last letters of X_t belongs to is enough to determine the quantity $Z_t := |X_{t+1}| - |X_t|$. Neglecting the case where X_t or X_{t+1} og-f length smaller than L, Z_t can thus be determined from a Markov chain on the state space $\{1,2\}$. The drift of X_t can then be easily deduced from the stationary measure of this Markov chain. This is the computation done in [21] for $L = 1$.

This argument remains valid both for the finite support case $L > 1$ and for colored random walks. Hence the computation can be made with or without the linearization trick.

Direct computation. Given $g \in G$, let $\underline{g}$, resp. $\overline{g}$ denote the first, resp. last letter of g. If $g = e, \underline{g} = \overline{g} = 1$. Suppose that at time t, $|X_t| > L$ so that neither in X_t nor X_{t+1} can equal e.

At time $t+1$, X_t is multiplied by some element g. Here are the different cases occuring. If $\underline{g}$ is not in the same group as $\overline{X_t}$, there is no cancellation when concatenating X_t and g, so $X_{t+1} = X_t g$ ends with the same letter as g, ie $\overline{X_{t+1}} = \overline{g}$. Otherwise there might be some cancellations and $\overline{X_{t+1}} = \overline{g}$ if and only if there is no complete cancellation between X_t and g. By complete cancellation we mean that g is exactly the inverse of the word formed by the $|g|$ last letters of X_t. If complete cancellation occurs, then $\overline{X_{t+1}}$ is bound not to be in the same group as $\overline{g}$.

Since there are only two groups, $\overline{g} = \underline{g}$ if and only if $|g|$ is odd. Thanks to the simple form of the transition probabilities in this model, one can deduce thus deduce that the probability the last letter of X_t goes from G_1 to G_2 is

$$Q_{1,2} = \sum_{n\geq 0} p_2(2n+1) + \sum_{n\geq 0} \frac{p_1(2n+1)}{k_1^{n+1}k_2^n} + \sum_{n\geq 1} p_1(2n)(1 - \frac{1}{k_1^n k_2^n}),$$

and the probability to stay in G_1

$$Q_{1,1} = \sum_{n\geq 2n+1} p_1(2n+1)(1 - \frac{1}{k_1^{n+1}k_2^n}) + \sum_{n\geq 1} \frac{p_1(2n)}{k_1^n k_2^n} + \sum_{n\geq 1} p_2(2n).$$

The sums above are finite as $p_i(n) = 0$ if $n > L$. By symmetry, exchange the roles of 1 and 2 to get $Q_{2,1}$ and $Q_{2,2}$. In the end, one obtains a 2×2 stochastic matrix Q which is the transition matrix of the Markov chain induced by the last letter of X_t in the sense that $\mathbb{P}\left[\overline{X_{t+1}} \in G_j \mid \overline{X_t} \in G_i, |X_t| > L\right] = Q_{i,j}$ for $i,j \in \{1,2\}$. Since X_t is transient it spends only a finite amount of time in the ball $\{g \in G, |g| \leq L\}$ so the ergodic theorem for Markov chains shows that the average amount of time $\overline{X_t}$ is in G_i converges to converges to $\pi(i)$, where π is the unique invariant measure of the transition matrix Q.

Now what is $Z_t := |X_{t+1}| - |X_t|$? Consider the cases previously identified. If $\underline{g}$ and $\overline{X_t}$ belong to different group, then $Z_t = |g|$. Otherwise, each cancellation occuring when concatenating X_t and g make the distance decrease by one. Hence in case of complete cancellation $Z_t = -|g|$. Otherwise, once the cancellations have been dealt with, the first step simply updates $\overline{X_t}$ and consequently does not increase the distance, while the other remaining steps increase it by one. Therefore the contribution of g is $-|g|$ in case of complete cancellation and $|g| - 1 - 2i$ otherwise, where i is the number of cancellations occurring when concatenating X_t and g.

Combining all these observations, the change in distance when going from group G_1 to group G_2 is

$$\tilde{Q}_{1,2} = \sum_{n\geq 0}(2n+1)p_2(2n+1) - \sum_{n\geq 0}(2n+1)\frac{p_1(2n+1)}{k_1^{n+1}k_2^n} + \sum_{n\geq 1} p_1(2n)\sum_{i=1}^{2n}(2n-1-2(i-1))\,\frac{1}{k_1\cdots k_{i-1}}\left(1-\frac{1}{k_i}\right) \quad (47)$$

and when staying in G_1:

$$\tilde{Q}_{1,1} = \sum_{n\geq 2n+1}\sum_{i=1}^{2n+1}(2n-2(i-1))p_1(2n+1)\frac{1}{k_1\cdots k_{i-1}}\left(1-\frac{1}{k_i}\right) - \sum_{n\geq 1}(2n)\frac{p_1(2n)}{k_1^n k_2^n} + \sum_{n\geq 1}(2n)p_2(2n). \quad (48)$$

As before, $\tilde{Q}_{2,1}, \tilde{Q}_{2,2}$ can be obtained by exchanging the roles of 1 and 2. In the end the drift of X_t is given by

$$\gamma = \pi(1)(\tilde{Q}_{1,1} + \tilde{Q}_{1,2}) + \pi(2)(\tilde{Q}_{2,1} + \tilde{Q}_{2,2}). \quad (49)$$

Computation with the linearization trick. Since we know specifically the random walk (X_t), we do not follow exactly the construction of the linearized random walk in Section 2 but adapt it to the example. What seems quite natural is the following: consider matrices $\tilde{p}_g = (\tilde{p}_g(i_1\cdots i_n))_{i_1\cdots i_n}$ indexed by reduced words $i_1 \neq i_2 \cdots \neq i_n$ of length $n \leq L-1$, including the case $n = 0$ of the empty word, written exceptionnally 0. For all $g \in S$ and all reduced word $i_1\cdots i_n$ of length $n \leq L-1$, set

$$\tilde{p}_g(0, i_1\cdots i_n) = \begin{cases} \frac{p_{i_0}(n+1)}{k_{i_0}} & \text{if } g \in G_{i_0} \text{ and } i_0 \neq i_1 \\ 0 & \text{otherwise} \end{cases}$$
$$\tilde{p}_g(i_1\cdots i_n, i_2\cdots i_n) = \begin{cases} \frac{1}{k_{i_1}} & \text{if } g \in G_{i_1} \\ 0 & \text{otherwise.} \end{cases} \quad (50)$$

It is easy to see that $\sum_{g\in S} \tilde{p}_g$ is a stochastic matrix. Thus the matrices $\tilde{p}_g$ can be used to define a colored random walk $(Y_t)_{t\geq 0}$. In Equation (50), the color 0 is to be interpreted as a neutral color, while $i_1 \cdots i_n$ is to be interpreted as the "remaining step" for Y_t, meaning that Y_t will be mutliplied by successive elements of $G_{i_1}, \dots, G_{i_n}$ before it arrives at the color 0. For instance the probability to go from $(e, 0)$ to $(g, 212)$ with $g \in G_1$ is $p_1(4)/k_1$. From $(g, 212)$, the walk then goes to some $(gh, 12)$, $h \in G_2$ with probability $1/k_2$, then to $(ghg', 1), g' \in G_1$ with probability $1/k_1$ and finally to $(ghg'h', 0)$ with probability $1/k_2$. Hence with probability $p_1(4)$ it moved between two states with color 0, from e to a uniform element of length 4 starting in group G_1. This should make clear that that Y_t linearizes the initial random walk X_t in the sense of Theorem 1: if Y_t starts at $(e, 0)$, then it behaves exactly like X_t when evaluated at the t-th return time at color 0.

Let us now apply the previous argument to this colored random walk. Let r be the number of colors used in the linearization and for $i = 1, 2$ set $\tilde{P}_i := \sum_{g\in G_i} \tilde{p}_g$. When Y_t is far from e, it induces a colored Markov chain on the state space $\{1, 2\} \times [r]$. Because the random walk is nearest-neighbor, we are spared of all the considerations necessary to deal with finite support. The transition matrix we are looking for is simply

$$Q = \begin{pmatrix} (1 - 1/k_1)P_1 & P_1/k_1 + P_2 \\ P_1 + P_2/k_2 & (1 - 1/k_2)P_2 \end{pmatrix} \tag{51}$$

which is the exact block analog of the matrix in [21].

Likewise the change in distance is simply given by the matrix:

$$\tilde{Q} = \begin{pmatrix} 0 & -P_1/k_1 + P_2 \\ P_1 - P_2/k_2 & 0 \end{pmatrix}. \tag{52}$$

The matrix Q defining an irreducible Markov chain, it admits a unique stationary measure π. As before, the ergodic theorem gives the drift of Y_t as $\sum_{v\in[r]} (\pi\tilde{Q})(v)$ and by Theorem 1 the drift of X_t is

$$\gamma = \left(\sum_{n=1}^{L} p_1(n) + p_2(n)\right) \sum_{v\in[r]} (\pi\tilde{Q})(v).$$

References

[1] Sébastien Blachère, Peter Haïssinsky, and Pierre Mathieu. Asymptotic entropy and Green speed for random walks on countable groups. *Ann. Probab.*, 36(3):1134–1152, 2008. 6

[2] Charles Bordenave and Hubert Lacoin. Cutoff at the entropic time for random walks on covered expander graphs. *arXiv preprint arXiv:1812.06769*, 2018. 10

[3] Philippe Bougerol and Jean Lacroix. *Products of random matrices with applications to Schrödinger operators*, volume 8. Springer Science & Business Media, 2012. 33, 36

[4] René Carmona and Jean Lacroix. *Spectral theory of random Schrödinger operators*. Springer Science & Business Media, 2012. 33, 36

[5] Yves Derriennic. Quelques applications du théoreme ergodique sous-additif. *Astérisque*, 74:183–201, 1980. 20

[6] Ken Dykema. Symmetric random walks on certain amalgamated free product groups. In *Topological and asymptotic aspects of group theory*, volume 394 of *Contemp. Math.*, pages 87–99. Amer. Math. Soc., Providence, RI, 2006. 7

[7] Lorenz Gilch. Asymptotic entropy of random walks on free products. *Electronic Journal of Probability*, 16:76–105, 2011. 7, 20

[8] Lorenz Gilch and François Ledrappier. Regularity of the drift and entropy of random walks on groups. *Publ. Mat. Urug.*, 14:147–158, 2013. 10

[9] Lorenz A. Gilch. Rate of escape of random walks on free products. *J. Aust. Math. Soc.*, 83(1):31–54, 2007. 7, 8, 20, 29

[10] Lorenz A. Gilch. Asymptotic entropy of random walks on free products. *Electron. J. Probab.*, 16:no. 3, 76–105, 2011. 8

[11] Sébastien Gouëzel. Analyticity of the entropy and the escape rate of random walks in hyperbolic groups. *Discrete Anal.*, pages Paper No. 7, 37, 2017. 10

[12] Robert H Haring-Smith. Groups and simple languages. *Transactions of the American Mathematical Society*, 279(1):337–356, 1983. 7

[13] J. William Helton, Tobias Mai, and Roland Speicher. Applications of realizations (aka linearizations) to free probability. *J. Funct. Anal.*, 274(1):1–79, 2018. 2, 5, 7

[14] Mairesse Jean. Random walks on groups and monoids with a Markovian harmonic measure. *Electronic Journal of Probability*, 10:1417–1441, 2005. 7, 10, 21, 24, 25, 26, 29

[15] Vadim A Kaimanovich and Anatoly M Vershik. Random walks on discrete groups: boundary and entropy. *The annals of probability*, pages 457–490, 1983. 6, 7, 20, 30, 32

[16] Steven P Lalley. Finite range random walk on free groups and homogeneous trees. *The Annals of Probability*, pages 2087–2130, 1993. 10, 33, 34, 35, 36

[17] François Ledrappier. Analyticity of the entropy for some random walks. *Groups Geom. Dyn.*, 6(2):317–333, 2012. 10

[18] François Ledrappier. Some asymptotic properties of random walks on free groups. *Topics in probability and Lie groups: boundary theory*, 28:117–152, 2001. 7, 10, 20, 22, 26, 28, 29, 32

[19] Franz Lehner. Computing norms of free operators with matrix coefficients. *American Journal of Mathematics*, 121(3):453–486, 1999. 7, 29

[20] Russell Lyons and Yuval Peres. *Probability on trees and networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 2017. 17, 19, 31

[21] Jean Mairesse and Frédéric Mathéus. Random walks on free products of cyclic groups. *Journal of the London Mathematical Society*, 75(1):47–66, 2007. 7, 20, 26, 29, 32, 38, 39, 41

[22] James A Mingo and Roland Speicher. *Free probability and random matrices*, volume 35. Springer, 2017. 2

[23] Tatiana Nagnibeda and Wolfgang Woess. Random walks on trees with finitely many cone types. *J. Theoret. Probab.*, 15(2):383–422, 2002. 8

[24] Eugene Seneta. *Non-negative matrices and Markov chains.* Springer Science & Business Media, 2006. 33

[25] Wolfgang Woess. *Random walks on infinite graphs and groups*, volume 138. Cambridge university press, 2000. 8